

CONFERENCE SERIES

Hans-Peter Schröcker, Manfred Husty (eds.)

**Proceedings of the 16th International
Conference on Geometry and Graphics**

Innsbruck, August 4-8, 2014

innsbruck university press

CONFERENCE SERIES



Hans-Peter Schröcker, Manfred Husty (eds.)

Proceedings of the 16th International Conference on Geometry and Graphics

Innsbruck, August 4-8, 2014

Hans-Peter Schröcker, Manfred Husty
Faculty of Engineering Science, Unit Geometry and CAD, Universität Innsbruck

© *innsbruck* university press, 2014

Universität Innsbruck

1st edition

All rights reserved.

www.uibk.ac.at/iup

ISBN 978-3-902936-46-2

Contents

On 3D Extension of the Simson-Wallace Theorem (P. Pech)	1
Aggregation of Polyhedron Modules on a Freeform Surface (P. Cutellic, et al.)	10
The Aid of GeoGebra to Teach and Learn Descriptive Geometry at the Faculty of Operation and Economics of Transport and Communications at the University of Žilina in Žilina, Slovakia (V. Čmelková)	18
The Analogue of Theorems Related To Wallace-Simson's Line in Quasi-Hyperbolic Plane (A. Sliepčević, I. Božić)	25
Analysis for Accessibility to the Convenient Facilities From the Residences (Ai Sakaki, R. Koba)	29
Analysis on Sequence of Architectural Space by Using VR Walk-Through System (K. Yasufuku)	38
Application of Curves and Surfaces of Higher Orders Obtained by Inversion in the Practice of Architecture (G. Đukanović, et al.)	45
Application of the Cross-Ratio to the Analysis of Architecture (R. Migliari, L. Baglioni)	54
An Approximation Method of Triangular Mesh Models for Generating Development Diagrams with Creases and Slits (K. Takahashi, T. Tachi, Y. Yamaguchi)	66
Arm-Z: A Modular Virtual Manipulator (M. Zawidzki, T. Nagakura) .	75
Assessment Method for an Edge Alignment Free Hybrid Image (P. Sriprian, Y. Yamaguchi)	81
Automated High Precision Texturing of 3D-Scans (M. Bornemann, S. Melzer, D. Lordick)	93
Automatic Range Scan Point Cloud Registration Using Hierarchical Levels and Feature Recognition Filters (A. Alonso, M.-H. Choi)	103
Bi-Arc Spirals in Minkowski Planes (G. Weiß, V. W. De Spinadel) . .	115

Brauner's Angle Formula and the Theory of Curves in Minkowski Planes and Spaces (E. Shonoda, G. Weiß)	121
Brunelleschi's Dome: An Overview Through 3D Digital Model About Geometrical Genesis and Proportional Theories (S. Bertacchi)	130
Caustic Surface and Quasifocal Line (A. Dvoretzky, T. Denysova)	142
Classification of Manifolds Resulting as Minkowski Operation Prod- ucts of Basic Geometric Point Sets (D. Velichová)	149
Coexistence of Geometry and Geology – Reconsidering Le Corbusier (M. Kato)	159
Color Scheme Scrapbook Using a Character Color Palette Template (R. Motegi, et al.)	167
Complexity of Public Spaces System Between Key Tall Buildings in City of Szczecin. Geometrical Aspect of Public Spaces in 3D City Model (A. Zwoliński)	175
Composite Concave Cupolae as Geometric and Architectural Forms (S. Mišić, M. Đ. Obradović, G. Đukanović)	187
Computing the Higher Dimensional Delaunay-Decomposition Using Depth-First-Search (M. Hünninger)	197
Congruent Stewart Gough Platforms With Non-Translational Self- Motions (G. Nawratil)	204
Contact of the Ruled Nondevelopable Surfaces (K. L. Panchuk, A. S. Ni- teyskiy)	216
Control Method of Combined Developable Surface Design by Affine Transformation and Locus (H. Suzuki)	224
Curvature Functions on a One-Sheeted Hyperboloid (B. Odehnal)	235
The Curved Horizon by Leonardo: Towards a New Perspective (D. Cal- isi)	245
Curves of Centres of Conic Pencils in Pseudo-Euclidean Plane (M. Katić Žlepalo)	257
Descriptive Geometry Education by Using Multimedia Tools (A. A. Cu- cakovic, N. K. Teofilovic, B. S. Jovic)	262
Design and Application Studies for a Cupola Forming Orbital Ar- rangement of Miura-Ori Basic Units (M. Barej, et al.)	266
Developing Spatial Ability for Quality Engineering Education (M. Ilić, M. Stavrić)	277

The Development of Modeling Tools for Pin-by-Pin Precise Reactor Simulations (Y. Ma, et al.)	285
Dynamic Geometry and Website Setup by Automatic Object Recognition of Free-Hand Drawings and Scans (K.-H. Brakhage)	291
The Educational Research Project GeodiKon: Pointing Accuracy, Strategies and Gender-Specific Effects (G. Maresch)	301
An Educational Trial on Drawing for True Shape by Using Auxiliary View Method Aided With CAD to the Students With Hearing Impairments (I. Honma, K. Kuwabara)	310
An Effectual Offset Algorithm for 3D Curve in Engineering Application (W. Liu, et al.)	314
Energetical Streams Visualization Using Interactive Sankey Diagrams (D. Ciminieri, et al.)	321
Enumeration of Deltahedral Graphs With up to 10 Vertices (N. Tsuruta, et al.)	328
Error Analysis of Problems on Surface Area of an Object Presented in a Three-View Orthographic Drawing as a Part of Mathematics Test Form (K. Shiina, S. Arai, T. Otsu)	336
A Eulogy of Ambiguity: Between Visual Perception and Quantum Mechanics (G. Caglioti, L. Cocchiarella)	343
Evaluation of Passing Performance of Self-Driven Particle Through Building (T. Matsumoto, K. Yasufuku, H. Abe)	355
Evaluation of Spatial Ability by Using a Silhouette of Solid Figure in Graphic Science Education (A. Takahashi, H. Abe)	363
Evaluation on Triangle-Triangle Intersection Tests Algorithms (H. Yu, Y. He, Y. Wang)	371
Existence and Uniqueness of Centers of Regular Polygons in Some Subclasses of IM-Quasigroups (S. Vidak)	375
Experiments With a Folding Multi-Agent System in the Design of Triangle Mesh Structures (B. Felbrich, et al.)	383
Exploring the Complexity. Digital Turn Towards Geometry in Contemporary Architecture and Urban Planning (P. Rubinowicz)	393
Eyesight Cartographies – Unfolding the Visual Sphere (J. V. Correia, et al.)	405
Fibonacci Triangles and Circle Chains and a Golden Bi-Arc Spiral in Non-Euclidean Planes (S. Mick, G. Weiß)	415

Foldable Truss-Z Module (M. Zawidzki, T. Nagakura)	423
Forming of the Spline Similar Linear Strip (D. S. Korchagin, K. L. Pan- chuk)	428
Freeform Surfaces Adaptation Using Planar Quadrilateral Facets (F. Gon- zález-Quintal, A. Sánchez-Parandiet, J. Barrallo)	437
The General Approach to the Solution of Typical Engineering Geo- metrical Problems (A. Y. Brailov)	444
Generalized Metallic Means Family (V. W. De Spinadel)	459
Generating the GROWTH Model Geometry Based on the Human Posture (S. Iwasawa, Y. Kawaguchi)	467
The Geometric and Computer Modeling Shaping Technical Surfaces (A. Lyashkov, V. Volkov)	475
Geometric Optimization in Minkowski Space (B. Blaschitz)	484
Geometric Perspective – Cylindrical Perspective (F. Doli)	495
Geometric Shapes That Generate Architectural Forms (C. Palestini)	505
The Geometrical Approach to Calculation of Boundary Layer Between Flows (O. Gumen, V. Mileikovskiy)	513
Geometrical Aspects of City Skyline – Tall Building Analysis (K. Czyńs- ka)	519
Geometry and Fuzzy Navigation System for Virtual Robot in the Drawing Environment (A. Sokas)	531
Geometry and Graphic Design (S. Chiarenza)	538
Geometry as a Tool for the Design of Military Architecture: The Expe- rience of Giovanni Battista Antonelli (S. Parrinello, S. Bertacchi)	548
The Geometry Behind the Octagonal Hall of Small Thermal Baths at Hadrian's Villa (L. Cipriani, F. Fantini)	560
Geometry of Gothic Vaults Based on Historical Sources (A. Kulig)	572
On the Geometry of Pseudo-Kossel and Bremsstrahlung Interferences (S. Enghardt, J. Bauch, F. Henschel)	580
The Geometry of the "Prospect Geometrique" by Micheli Du Crest (1754) – a Quantitative Analysis Within a Hybrid Least- Squares Adjustment Framework (H. Kager, M. Rickenbacher, A. Roncat)	590
The Graphic Interpretation of the Co-Occurrence Network Diagram in Different Writing Styles (M. Yamade, H. Abe)	600

Hexahedral Mesh Generation Using Adaptive Sampling Based on Sharp Features (Y. Imai, H. Hiraoka, H. Kawaharada)	607
Historical Perspectives of Perspective (A. H. Tan, F. H. Tan, F. M. Croft Jr.)	614
How to Promote Student Creativity and Learning Using Tutorials in Teaching Graphics and Visualisation (A. Hast)	626
Hull Lines Plan. History. Comparison of Methodologies for 3D Model Generation (F. Fadon, et al.)	634
HyperCAL3D: A Computer Application to Support the Teaching and Learning of Descriptive Geometry (F. G. Teixeira, S. L. Dos Santos)	643
Image Processing Education in the Department of Information Engineering, College of Industrial Technology – Utilization of Spreadsheet Software and Web System (M. Mori, et al.)	655
Imagined Space, Real Space. The Curtain of the Theatre and the Central Hall of the Royal Palace of Caserta (N. Pisacane)	661
Improvement of Quadric Error Metric Mesh Simplification by Reeb-Graph-Based Topological Information (C. Devahastin Na Ayudhya, P. Kanongchaiyos)	672
The Incircle Hyperboloid and Ellipsoid of the Conjugate Pavillet Tetrahedra (A. Pavillet)	680
Interaction Among Courses (J. Beban-Brkić, M. Šimić Horvath)	690
Introducing 3D Modeling Into Geometry Education at Two Technical Faculties at the University of Zagreb (S. Gorjanc, H. Halas, E. Jurkin)	697
Intriguing Intersections and Useful Unfoldings (G. Barczik)	706
Iterative Form Finding for the Layout of Irregular Reciprocal Framework Structures (T. Kuhlmann, D. Lordick)	712
Kinematic Exploration of 1-Dof Origami Mechanisms (Y. Klett, P. Misdendorf)	719
Kinematics of a Mobile Overconstrained RRRCR Loop (J. J. Cervantes-Sánchez, P. Zsombor-Murray)	726
Learn by Eating – An Easy Way to Approach Geometrical and Mathematical Aspects (E. Marchetti, L. Rossi-Costa)	732
Line Inversion in the Quasi-Hyperbolic Plane (H. Halas, N. Kovačević, A. Sliepčević)	739

The Locus-Function of the Perspective Circle (M. Sejfried)	749
Logic (Systematic) Vs. Automation (Mechanisms) (A. Mollicone) . .	752
Logical-Constructive Teaching Approach to the Theme “Creation of Surfaces” (N. Kaygorodtseva, T. Kaygorodtseva)	758
Map Projection of the World Map by Leonardo Da Vinci (M. Lapaine)	764
On Method of Producing a Deformed Small Figure Based on an Actual Person’s Body (S. Cho, H. Sato, H. Yoshii)	774
A Method Study of Rapidly Making Large-Scale Topographic Map (A. Jia, N. Wang, P. Ruan)	778
Methodological Bases of Geometric Modeling Multifactorial Pro- cesses (V. Volkov, O. Ilyasova, M. Chizhik)	784
Modeling of Ship Hulls With NURBS Curves and Surfaces (F. Perez)	794
Modern Book on Geometry (V. Shelomovskii, S. Nosulya)	802
The Monge Three Point Space Resection Problem (R. Migliari, F. Fallavol- lita, M. Salvatore)	809
Multivariant Data in Complex Quiz Layout for Computer Engineering Graphics (D. Makuteniene)	820
Navigating Multi-Dimensional Landscapes in Foggy Weather as an Analogy for Generic Problem Solving (D. Rutten)	826
Navigation Functions of Pictorial Schematics During the Edo Period With Emphasis on Travel Maps (T. Konishi, S. Kaneko) . . .	840
A New Condition of Perpendicularity in Isometrical Cavalier Projec- tion (V. Cardone, B. Messina)	847
A New Focus for Teaching and Learning With Descriptive Geometry’s Tutorials Through the Web (G. Valencia García)	858
A New Triangle Theorem to Solve the Inverse Kinematics Problem for Characters With Highly Articulated Limbs (E. A. Rosales, L. E. Falcón-Morales)	865
Null Polarities as Generators of the Projective Group (D. Klawitter)	873
An Open Source Work Flow: A Digital Building Reconstruction and its Presentation (A. Mazzucchi)	885
Optimisation of Bundled Routes (V. Parque, M. Kobayashi, M. Higashi)	893
Overconstrained Mechanisms Based on Special Planar Chains (O. Rö- schel)	903
The Parabola in Universal Hyperbolic Geometry II: Canonical Points and The Y-Conic (A. Alkhalidi, N. Wildberger)	912

Parametric Approach to Data Interpolating Curves Using the Sum of Bell-Shaped Functions (I. Taras)	922
Patterns on Triply Periodic Polyhedra (D. Dunham)	928
Perspectival Geometries in the Sacred Mount of Ossuccio (G. Amoruso, A. Sdegno, S. Masserano)	936
Perspective Between Fiction and Function: Pattern Mutations Through Science and Art (L. Cocchiarella)	944
Perspective Concepts. Exploring Seeing and Representation of Space (C. Leopold)	956
Perspective: Theories and Experiments on the “Veduta Vincolata” (Restricted Sight) (R. Migliari, J. Romor)	968
Point in Polygon Via Epi-Hypo Graphs, Homotopy and Hopf’s Degree Theorem (S. Sinha, et al.)	982
Qualitative Investigation on Formative Design Process Using 3D CAD by Industrial Designers (M. Nishii, T. Saito)	1002
Real Life and Real Representation – Educating Along and Across the Boundaries of Geometry and Graphics (N.-C. Fritsche)	1012
A Recyclable Möbius Strip (M. Pfurner, J. Schadlbauer, H.-P. Schröcker)	1024
The Representation on the Cylindrical Rotary Surface – Inverse Panorama (J. Dzwierzynska)	1031
Research and Practice of Engineering Design Training With a Tight In- tegration Between Design and Engineering Graphics (Z. Ning- rong, X. Qinghua)	1037
Research of Applications on Synchronous Modeling Technology (L. Wang, et al.)	1043
A Research of Multimedia Teaching Materials for 3-Dimension CAD Education (H. Tominaga, et al.)	1048
The Research on Graph, Graphics and Graphics Science (B. Han, et al.)	1055
Research on the Test Sheet Organization for Engineering Graphics (L. Guan, et al.)	1061
Rigid Quadrilateral Folding Strategies for Surface Design in Archi- tecture (A. Wiltsche, M. Stavrić)	1066
On the Role of Circular Sections of Quadric Surfaces. The Elaboration of the Topic by Two Creative Geometric Student’s Tasks (A. A. Cucakovic, et al.)	1075

The Role of Hand-Drawing in the Representation of Architecture and Landscape (M. P. Marabotto)	1082
Ruled Surfaces of Constant Slope in 3-Minkowski Space (Ž. Milin-Šipuš, L. Primorac Gajčić)	1087
Sejfriedian: Existence, Uniqueness, Constructing and the Proof of Properties (V. Shelomovskii)	1095
A Self Checking CAD Tool for Mechanical Drawings in Introductory Courses (I. Tanaka)	1111
Self-Repeating Trajectories of Light Rays (M. Manevich, E. Itskovich, N. Shvalb)	1117
Several Sources of Shapes (S. Filipowski)	1125
Simulating the Construction of China's Song-Style Dougong Using Digital Graphics (S. Hao, et al.)	1130
Simulating the Construction Process of the Roman Colosseum Using Digital Graphics (A. H. Tan, F. M. Croft Jr., F. H. Tan)	1142
Some Cases Using Descriptive Geometry in Various Courses (S. Nagashima)	1154
Space Description Problems in the Secondary School (Their Possible Causes and Problem Solving) (R. Kiss-György)	1159
On the Special Surfaces Through the Absolute Conic With a Singular Point of the Highest Order (S. Gorjanc, E. Jurkin)	1168
Students' Ability to Model Parts From Various Types of Drawings and the Relationship to Other Measures in Engineering Graphics Courses (Th. Branoff, M. Dobelis)	1174
A Study on Morphological Interpretation of the Façade Design and Form Design in Architecture (K. Wada, H. Abe)	1184
A Study on Spatial Cycloid Gearing (H. Stachel, G. Figliolini, J. Angeles)	1192
A Study on the Parametric Architecture – Development of Computer Graphic Materials for Architectural Design Education (N. Ando, et al.)	1206
Study on Three-Dimensional Visualization of the Cable-Stayed Bridge (Z. Wang, C. Hu, M. Li)	1218
SunSys – A Case Study on Geometric Complexity in Computational Design (M. Hemmerling, J. Böke)	1222
A Survey on the Spatial Abilities of Pre-University Students – On Scores of Mental Cutting Tests (E. Tsutsumi, et al.)	1226

Theory and Practice in the Implementation of Illusionistic Ceiling Painting at Palazzo Moroni in Bergamo Italy (G. Mele, G. Burratti, F. Rovo)	1236
Thirteen Years of MRT – Results, Options and Dilemmas (D. Kusar, M. Volgemut)	1248
The Three-Dimensional Modeling and the Finite Element Analysis of the Involute Hyperbolic Arch Dam (T. Du, Q.-Q. Mao, Q. Du)	1257
Traditional Descriptive Geometry Education in 3D-CAD/CG Era (K. Suzuki)	1264
Translation Surfaces and Applications (L. Pletenac, Ž. Milin-Šipuš) .	1272
Trullis – Architectural Archetypes (G. Colagreco, H. Ortale, G. Weiß)	1283
Two Possible Variants of Geometrical Model of Constructing Kinematic Surfaces on the Base of Interior Revolving One Axoid by Another One (G. S. Rachkovskaya, Y. N. Kharabayev, N. S. Rachkovskaya)	1291
Use of Geometry at Creation or the Analysis of Design and Art Objects (I. Kuznetsova)	1295
The Use of Multi-Media in the Teaching of Educational Unit for the Methods of Altering Some Patterns of Women’s Clothing to Overcome the Problem of Fitting (T. Alsarhan, H. Abdelfattah)	1298
The Use of Photographs in the Teaching/Learning of Descriptive Geometry (B. Kotarska-Lewandowska)	1310
Using Computer Technologies for Educational Process on Graphic Subjects at MPEI (E. Kasatkina, Y. Stepanov, V. Kaurkin) . .	1316
Using Optical Illusion Patterns Affixed to Toy Blocks for Learning Human Errors in Three-Dimensional Projection (T. Ohtani, K. Maruya)	1320
Visualising Variability: Confidence Regions in Level Set Estimation (H. Jankowski, L. Stanberry)	1328
Visualization of Travel Time by Transportation Service With Concentric Circles (M. Shiraishi, A. Inoue, M. Shinya)	1340
Visualization With Visibility of Higher Dimensional and Non-Euclidean Geometries (J. Katona, et al.)	1345
Visualizing Normal Equations in Least-Squares Adjustment (A. Roncat)	1355
The Working Drawing of Architecture: A Teaching Experience of Reflection on Codes of Representation (G. Savarro)	1362

Author Index	1373
------------------------	------

ON 3D EXTENSION OF THE SIMSON–WALLACE THEOREM

Pavel PECH

University of South Bohemia, Czech Republic

ABSTRACT: The well-known Simson–Wallace theorem reads: If P is a point in the circumcircle of a triangle ABC then orthogonal projections of P onto the sides of ABC are collinear.

This theorem has several generalizations. Generalization which is ascribed to J. D. Gergonne states: If P is a point of a circle which is concentric with the circumcircle of a triangle ABC then orthogonal projections of P onto the sides of ABC form a triangle with constant area.

3D extension of the Simson–Wallace theorem is as follows: Let K, L, M, N be orthogonal projections of a point P to the faces BCD, ACD, ABD, ABC of a tetrahedron $ABCD$. Then the point P such that the tetrahedron $KLMN$ has constant volume belongs to a cubic surface (1). This theorem is proved by the method which differs from the method used in the past (Theorem 1). Next the main theorem (Theorem 2) which is converse to the Theorem 1 is proved.

Furthermore we'll verify the Theorem 2 for a regular tetrahedron by the method of descriptive geometry and the use of dynamic geometry software. To do this we take advantage of the fact that this cubic surface can be represented by a parametric system of conics which lie in mutually parallel planes (Theorem 3).

Keywords: Simson–Wallace theorem, loci, cubic surface, Cayley cubic, Monge projection

1. INTRODUCTION

The well-known Simson–Wallace theorem belongs to the most beautiful theorems of plane geometry. The Simson–Wallace theorem [4] reads:

If P is a point in the circumcircle of a triangle ABC then orthogonal projections of P onto the sides of ABC are collinear.

There exist several generalizations [6, 8, 10–12] of this theorem. Generalization which is ascribed to Gergonne [2] is as follows, Fig. 1:

If P is a point of a circle which is concentric with the circumcircle of a triangle ABC then orthogonal projections of P onto the sides of ABC form a triangle with constant area.

If the area is zero we get the classical Simson–Wallace theorem.

One could think that a generalization of the Simson–Wallace theorem on a tetrahedron into three dimensional space leads, by analogy with

a planar case, to a sphere. But this is not the case. In [10–12] the 3D extension of the Simson–Wallace theorem on a tetrahedron $ABCD$ is given:

Theorem 1: *Let K, L, M, N be orthogonal projections of an arbitrary point P to the faces BCD, ACD, ABD, ABC of a tetrahedron $ABCD$. Then the point P such that the tetrahedron $KLMN$ has constant volume s belongs to the cubic surface (1).*

First we'll give the proof of the Theorem 1, so that we could realize the method we use throughout the paper. This proof differs from the proof in [11, 12], where the software CoCoA [1] was used. To avoid successive elimination here the software Epsilon [14] was applied. All computational methods we use in this paper are the methods of the theory of automated geometry theorem proving [2, 3, 13, 14].

In the Theorem 2 we'll show that also the converse theorem to the Theorem 1 holds. This en-

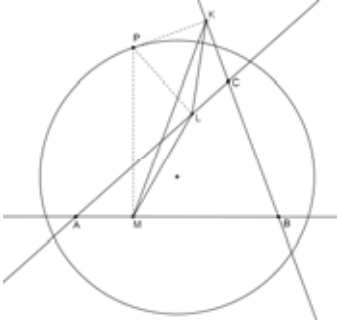


Figure 1: If P moves along the circle the triangle KLM has constant area

tures that the locus of the point P is the "whole" surface (1). The author tried to prove it a few years ago without success [12]. It turns out that the use of the software Epsilon which is based on Wu–Ritt method [14, 15] can help in solving this problem.

Further some properties of the cubic surface (1) are investigated. If the volume s equals zero then we obtain a special cubic surface called the Cayley cubic. In the Theorem 3 it is shown that if we take a regular tetrahedron and $s = 0$ then its associate Cayley cubic can be represented by a pencil of conics which are parallel to two opposite edges of the tetrahedron.

In the last section we verify the Theorem 2 by the methods of descriptive geometry together with dynamic geometry system GeoGebra [5].

2. THE MAIN THEOREM

First we recall the basic ideas of the proof of the Theorem 1. The novelty of this proof is the use of the software Epsilon with characteristic sets whereas in the previous proof [11, 12] we used the software CoCoA [1] and successive elimination of variables. Secondly we'll prove the main theorem — Theorem 2 — which states that also the converse statement holds.

Proof of Theorem 1: Let K, L, M, N be orthogonal projections of P onto the faces BCD, ACD, ABD

and ABC of a tetrahedron $ABCD$.

Choose a rectangular system of coordinates so that $A = (0, 0, 0)$, $B = (a, 0, 0)$, $C = (b, c, 0)$, $D = (d, e, f)$, $K = (k_1, k_2, k_3)$, $L = (l_1, l_2, l_3)$, $M = (m_1, m_2, m_3)$, $N = (n_1, n_2, n_3)$, $P = (p, q, r)$ and let s be oriented volume of $KLMN$. Suppose that $a \neq 0$, $c \neq 0$, $f \neq 0$ since otherwise $ABCD$ is planar. Then:

$$PK \perp BCD \Leftrightarrow$$

$$h_1 : (b-a)(p-k_1) + c(q-k_2) = 0,$$

$$h_2 : (d-a)(p-k_1) + e(q-k_2) + f(r-k_3) = 0,$$

$$K \in BCD \Leftrightarrow$$

$$h_3 : -acf - aek_3 + afk_2 + ack_3 + cfk_1 + bek_3 - cdk_3 - bfk_2 = 0,$$

$$PL \perp ACD \Leftrightarrow$$

$$h_4 : b(p-l_1) + c(q-l_2) = 0,$$

$$h_5 : d(p-l_1) + e(q-l_2) + f(r-l_3) = 0,$$

$$L \in ACD \Leftrightarrow h_6 : cfl_1 + bel_3 - cdl_3 - bfl_2 = 0,$$

$$PM \perp ABD \Leftrightarrow h_7 : p-m_1 = 0,$$

$$h_8 : d(p-m_1) + e(q-m_2) + f(r-m_3) = 0,$$

$$M \in ABD \Leftrightarrow h_9 : em_3 - fm_2 = 0,$$

$$PN \perp ABC \Leftrightarrow$$

$$h_{10} : p-n_1 = 0, \quad h_{11} : b(p-n_1) + c(q-n_2) = 0,$$

$$N \in ABC \Leftrightarrow h_{12} : n_3 = 0.$$

Conclusion h_{13} is of the form:

$$\text{Volume of } KLMN = s \Leftrightarrow$$

$$h_{13} : \begin{vmatrix} k_1 & k_2 & k_3 & 1 \\ l_1 & l_2 & l_3 & 1 \\ m_1 & m_2 & m_3 & 1 \\ n_1 & n_2 & n_3 & 1 \end{vmatrix} - 6s = 0.$$

We eliminate dependent variables $k_1, k_2, k_3, l_1, l_2, l_3, m_1, m_2, m_3, n_1, n_2, n_3$ in the system of algebraic equations $h_1 = 0, h_2 = 0, \dots, h_{13} = 0$ with variable ordering $p \prec q \prec r \prec a \prec b \prec \dots \prec f \prec s \prec k_1 \prec k_2 \prec \dots \prec n_3$. The use of the Wu–Ritt method with characteristic sets in Maple package Epsilon [14, 15] gives

with(epsilon);

CharSet({h1, h2, h3, h4, h5, h6, h7, h8, h9, h10, h11, h12}, [p, q, r, a, b, c, d, e,

$f, s, k[1], k[2], k[3], l[1], l[2], l[3], m[1], m[2], m[3], n[1], n[2]]$;

the equation

$$F := ac^2f^3G + sQ = 0, \quad (1)$$

where

$$G = c^2f^2p^2q + cf(e^2 + f^2 - ce)p^2r + cf^2(a - 2b)pq^2 + cf^2(a - 2d)pr^2 + 2cef(b - d)pqr + b(b - a)f^2q^3 + f(be(a - b) + cd(d - a) + cf^2)q^2r + f^2(b^2 - ab + c^2 - 2ce)qr^2 + (be(a - b) + cd(d - a) + ce(e - c))fr^3 - ac^2f^2pq + acf(ce - e^2 - f^2)pr + abcf^2q^2 + (a(c^2d - 2bce + be^2) - (cd - be)^2 + f^2(ab - b^2 - c^2))fqr + (ce^2(ab + ad - 2bd) + c^2de(d - a) + be^3(b - a) + f^2(acd - be) + e(b^2 + c^2))r^2$$

and

$$Q = 6(e^2 + f^2)((cd - be)^2 + f^2(b^2 + c^2))((c(a - d) - e(a - b))^2 + f^2((a - b)^2 + c^2)).$$

As $a \neq 0$, $c \neq 0$ and $f \neq 0$ we see that (1) describes a cubic surface. \square

Further we'll show that the converse theorem holds as well.

We'll prove that for every point $P = [p, q, r]$ of the surface (1), the points K, L, M, N are complanar. We want to show that from the conditions $h_1 = 0$, $h_2 = 0$, ..., $h_{12} = 0$ and $F = 0$ the conclusion $h_{13} = 0$ follows. We enter

```
with(epsilon);
Simson:=Theorem({h1,h2,h3,h4,h5,
h6,h7,h8,h9,h10,h11,h12,F},{h13},
[f,e,d,c,b,a,p,q,r,s,k[1],k[2],
k[3],l[1],l[2],l[3],m[1],m[2],
m[3],n[1],n[2]]):
Prove(Simson);
```

and get the answer

The theorem is true under the following subsidiary conditions:

- 1) $c \neq 0$,
- 2) $f \neq 0$,

$$3) e^2 + f^2 \neq 0,$$

$$4) (be - cd)^2 + f^2(b^2 + c^2) \neq 0,$$

$$5) (e(a - b) - c(a - d))^2 + f^2((a - b)^2 + c^2) \neq 0.$$

All conditions 1) – 5) are fulfilled as we suppose that $c \neq 0$ and $f \neq 0$.

Note the ordering of variables with variable ordering $f \prec e \prec d \prec c \prec b \prec a \prec p \prec q \prec r \prec s \prec k_1 \prec k_2 \prec \dots \prec n_3$. By this ordering we get subsidiary conditions in such a form which directly depends on the non-zero values c and f . We can accept them without restricting hypotheses of the theorem. We can state the main theorem:

Theorem 2: *Let K, L, M, N be orthogonal projections of a point P to faces BCD, ACD, ABD, ABC of a tetrahedron $ABCD$. Then the locus of P such that the tetrahedron $KLMN$ has constant volume s is the surface (1).*

We'll call the surface F the associate surface of the tetrahedron $ABCD$ and the tetrahedron $ABCD$ the associate tetrahedron of the surface F .

3. PROPERTIES OF THE SURFACE F

The surface (1) belongs to the family of algebraic surfaces of the third degree called cubic surfaces, or briefly just cubics [7], [9]. We'll mention only those properties of (1) that we need in this article.

First suppose that $s = 0$. This surface has four singular points, the maximum number of singular points of a general cubic surface, which are placed in the vertices of the tetrahedron $ABCD$. Cubic surface with four singular points is called the *Cayley cubic* (do not change it with Cayley ruled surface). Cayley cubic contains 6 lines AB, BC, CD, DA, AC and BD , the edges of $ABCD$ which are torsal lines of (1). Another 3 straight lines of the surface are intersections of tangent planes along opposite edges of $ABCD$. These three lines are complanar [11], [12].

If $s \neq 0$ then the surface (1) does not have any singular points and the vertices A, B, C, D of the associate tetrahedron do not lie on it. Note that the equation of this surface differs from the previous Cayley cubic only in a nonzero constant s .

In the following we'll demonstrate how the cubic surface (1) changes its shape with respect to dif-



Figure 2: Cayley cubic $p^2q + pq^2 + p^2r + q^2r + pr^2 + qr^2 - pq - pr - qr = 0$ with four singular points at vertices of the tetrahedron

ferent values of the volume s .

Let the basic tetrahedron $ABCD$ be given with its vertices $A = (0, 0, 0)$, $B = (1, 0, 0)$, $C = (0, 1, 0)$ and $D = (0, 0, 1)$. Setting the values $a = 1$, $b = 0$, $c = 1$, $d = 0$, $e = 0$ and $f = 1$ into (1) we get the equation of the associate cubic surface

$$p^2q + pq^2 + p^2r + q^2r + pr^2 + qr^2 - pq - pr - qr + 18s = 0.$$

For $s = 0$ we get

$$p^2q + pq^2 + p^2r + q^2r + pr^2 + qr^2 - pq - pr - qr = 0$$

which is the Cayley cubic with singular points at the vertices A, B, C and D , Fig. 2.

Now let us consider the same tetrahedron $ABCD$ as in the previous case with non zero value of the volume s of the related tetrahedron K, L, M, N .

For $s > 0$, for instance for $s = 1/1800$, we get the following cubic surface Fig. 3. Note that we obtained the cubic surface which consists of two separated parts.

Similarly the Cayley surface with negative s , for instance for $s = -1/1800$, is as follows, Fig. 4.



Figure 3: Cubic surface $p^2q + pq^2 + p^2r + q^2r + pr^2 + qr^2 - pq - pr - qr + 1/100 = 0$ with positive s

Thus we obtain three types of cubic surfaces in accordance with the sign of the volume s , cf. the Gergonne generalization of the Simson–Wallace theorem in a plane, where we also get three types of loci (circles) [11].



Figure 4: Cubic surface $p^2q + pq^2 + p^2r + q^2r + pr^2 + qr^2 - pq - pr - qr - 1/100 = 0$ with negative s

4. REGULAR TETRAHEDRON AND ITS ASSOCIATE SURFACE

In this section we'll study a regular tetrahedron and its associate cubic surface.

Putting $A = (0, 0, 0)$, $B = (2, 0, 0)$, $C = (1, \sqrt{3}, 0)$ and $D = (1, 1/\sqrt{3}, \sqrt{8/3})$ we get a regular tetrahedron. The cubic surface associated with this tetrahedron for a given s has the equation

$$3p^2q + \frac{3}{2}\sqrt{2}p^2r - q^3 + \frac{3}{2}\sqrt{2}q^2r - \sqrt{2}r^3 + 2\sqrt{3}q^2 + \frac{5}{2}\sqrt{3}r^2 - 6pq - 3\sqrt{2}pr - \sqrt{6}qr + \frac{243}{16}\sqrt{6}s = 0.$$

See this surface for $s = 0$ in Fig. 5.

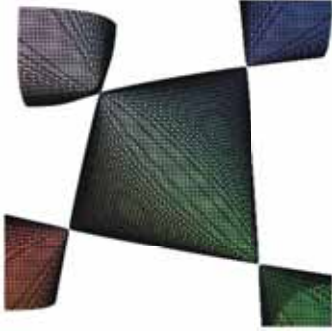


Figure 5: Cayley cubic associated with regular tetrahedron

To avoid radicals and to obtain a simpler equation of a Cayley cubic associated with a regular tetrahedron, we'll use the mapping

$$(p, q, r, 1) \rightarrow (p, q, r, 1) \cdot M, \quad (2)$$

where

$$M = \begin{pmatrix} -1 & -\frac{1}{\sqrt{3}} & \sqrt{\frac{2}{3}} & 0 \\ 1 & -\frac{1}{\sqrt{3}} & \sqrt{\frac{2}{3}} & 0 \\ 0 & \frac{2}{\sqrt{3}} & \sqrt{\frac{2}{3}} & 0 \\ 1 & \frac{1}{\sqrt{3}} & -\sqrt{\frac{2}{3}} & 1 \end{pmatrix}$$

which maps the regular tetrahedron $ABCD$ with vertices $A = (0, 0, 0)$, $B = (2, 0, 0)$, $C = (1, \sqrt{3}, 0)$ and $D = (1, 1/\sqrt{3}, \sqrt{8/3})$ on the regular tetrahedron $A'B'C'D'$ with vertices $A' = (1, 0, 0)$, $B' =$

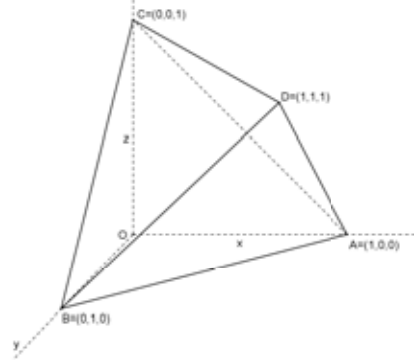


Figure 6: Regular tetrahedron in another coordinate system

$(0, 1, 0)$, $C' = (0, 0, 1)$ and $D' = (1, 1, 1)$, Fig. 6. The associate cubic surface for a given s transforms into

$$4pqr - (p + q + r - 1)^2 - \frac{81}{4}s = 0. \quad (3)$$

Similarly for a regular tetrahedron $ABCD$ with vertices $A = (a, 0, 0)$, $B = (0, a, 0)$, $C = (0, 0, a)$ and $D = (a, a, a)$ and s we get by (2) the equation of an associate cubic surface in the form

$$4pqr - a(p + q + r - a)^2 - \frac{81}{4}s = 0. \quad (4)$$

Note that the constant by s in (4) does not depend on a .

Now we'll investigate properties of the surface (4) for $s = 0$

$$4xyz - a(x + y + z - a)^2 = 0. \quad (5)$$

We changed the notation from p, q, r to common x, y, z which is now more comfortable. Suppose that $a > 0$.

Exploring sections of (5) with a pencil of mutually parallel planes $z = k$, where k is a real parameter, we get

$$4kxy - a(x + y + k - a)^2 = 0.$$

This yields for a fixed parameter k the conic

$$ax^2 + 2(a - 2k)xy + ay^2 + 2a(k - a)x + 2a(k - a)y + a(k - a)^2 = 0 \quad (6)$$

with a canonical form

$$2kx^2 + 2(a - k)y^2 - ak(a - k) = 0. \quad (7)$$

Exploring (7) we get: If $k < 0$ or $k > a$ the conic is a hyperbola, if $k = 0$ or $k = a$ the conic is a double line, and finally if $0 < k < a$ the conic is an ellipse, Fig. 7.

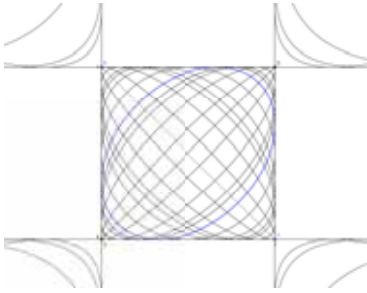


Figure 7: Horizontal view of planar sections of a Cayley cubic — the system of conics

We can also study planar sections of (5) through a line which lies on the cubic surface. Recall that every plane intersects a cubic surface in a cubic curve.

If we want to get a conic section then a related cubic curve must be decomposed into a conic and a line. Thus the only conics on a cubic surface we get if the planar sections pass through a line which belongs to the cubic surface. As a cubic surface contains 27 lines (with multiplicities, some of them may be imaginär or lie at infinity), then we get 27 pencils of conics lying on the cubic.

In the case of a regular tetrahedron $ABCD$ the tangent planes of its associate Cayley cubic along a pair of opposite edges are mutually parallel and intersect at a line at infinity. This implies that planar sections of (5) which are parallel to two

opposite edges of a regular $ABCD$ must be conics.

Hence for the surface (5), which is associated with a regular tetrahedron, we obtain the following theorem:

Theorem 3: *The Cayley cubic (5) associated with a regular tetrahedron can be expressed as a pencil of conics (6) which lie in planes which are parallel to two opposite edges of a tetrahedron.*

In the next section we'll apply this theorem to numerical verification of the Theorem 2.

5. VERIFICATION OF THE MAIN THEOREM IN DGS

We have seen that the 3D extension of Simson–Wallace theorem leads to a cubic surface and not to a sphere as one could expect. Whereas in a planar case of the Simson–Wallace theorem we can easily verify that the locus is the circumcircle of a triangle either by the ruler and compasses or by DGS, in 3D case the verification is problematic. Perhaps the following could help in 3D (numerical) verification by 3D software DGS of statements in which cubics (or quadrics) occur.

In this section we'll demonstrate the verification of the Theorem 2 for a regular tetrahedron by the method of descriptive geometry together with the

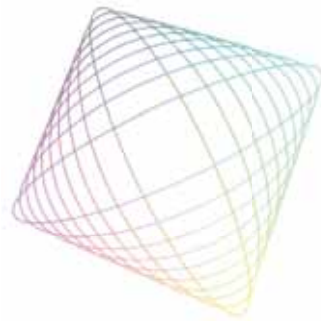


Figure 8: Cayley cubic associated with regular tetrahedron — the tetrahedral part

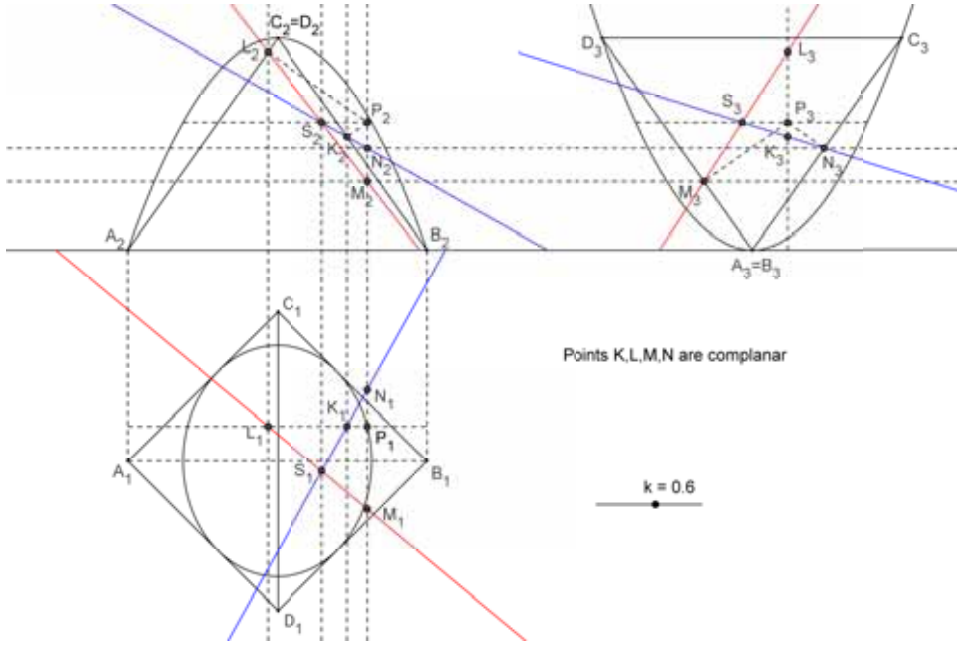


Figure 9: Points K, L, M, N are complanar

use of dynamic geometry software GeoGebra [5]. We will aim at the most interesting part of (5) — the tetrahedral part — or more precisely, the part which is described by the system (6) of conics for $0 \leq k \leq 1$, Fig. 8.

Verification of the Simson–Wallace theorem in dynamic geometry system will be performed using the Monge orthogonal projection onto two mutually orthogonal planes. We'll prove that the feet K, L, M, N of perpendiculars from a point P of the Cayley cubic which is associated with a regular tetrahedron $ABCD$ onto the faces of $ABCD$ are complanar. To do this it suffices to show that the straight lines KL and MN intersect at the point S , Fig. 9.

Denote horizontal projections with the index 1 whereas vertical and side projections with the indices 2 and 3.

We place the tetrahedron in a such position that the edge AB lies in the horizontal projection plane and is parallel to the vertical projection plane. The opposite edge CD is parallel to the horizontal plane and orthogonal to the vertical projection plane. To obtain a better view we rotate the surface (5) by an angle 45° .

In the position in the Fig. 9 it holds

$$A = (-\frac{1}{\sqrt{2}}, 0, 0), B = (\frac{1}{\sqrt{2}}, 0, 0), C = (0, -\frac{1}{\sqrt{2}}, 1) \text{ and } D = (0, \frac{1}{\sqrt{2}}, 1).$$

The associate Cayley cubic has the equation

$$2x^2z - 2y^2z + 2y^2 + z^2 - z = 0. \quad (8)$$

By the Theorem 3 this Cayley cubic may be represented by the pencil of conics

$$2kx^2 + 2(1-k)y^2 - k(1-k) = 0. \quad (9)$$

which lie in the planes $z = k$.

Note that (9) we get from (7) for $a = 1$.

Let a point P move along an ellipse of the surface (8). From the point P we construct projections of the feet K, L of perpendiculars to the faces BCD and ACD . We used the side projection to simplify the construction. Similarly, we construct projections of the feet M, N of perpendiculars to the faces ABC and ABD . Note that the vertical projection of the associate Cayley cubic (8) of $ABCD$ is a parabola $z = -2x^2 + 1$ and a line $z = 0$.

The whole system is fully interactive. We can move both an ellipse of a section using a slider with a parameter k of the ellipse (9) and a point P along the ellipse using a pointer. In this way we can cover the whole tetrahedral part. Finally we construct the intersection S_1 of the lines K_1N_1 and L_1M_1 , and the intersection S_2 of lines K_2N_2 and L_2M_2 . Now using the window "Relation between two objects" we get the answer that S_2 lies on the line through S_1 and which is orthogonal to x axis, i.e., the lines KN and LM are concurrent. This can be connected with the text *Points K, L, M, N are complanar*. Thus moving the point P along the Cayley cubic, we see the text that the points K, L, M, N are complanar. If we detach the point P from the cubic, i.e. from the ellipse, using the window "Detach point", the text disappears. In this way we are able to verify the validity of the theorem for infinitely many positions of the point P . The verification by DGS is now complete.

Note that this verification is based on numerical description of geometry objects. From the mathematical point of view this can not be considered as a strict mathematical proof.

6. CONCLUSIONS

The Wallace–Simson theorem has been generalized several times in the history. The 3D extension given in this paper is based on results of commutative algebra in the last third of the last century.

In the paper we proved that if K, L, M, N are orthogonal projections of a point P onto the faces of a tetrahedron $ABCD$ then the locus of the point P such that the tetrahedron $KLMN$ has a given volume s is the cubic surface (1).

A special case of the cubic surface (1) for $s = 0$ is a Cayley cubic. We have shown that the Cayley cubic which is associated with a regular tetrahedron can be represented by a pencil of conics which lie in planes parallel to the opposite edges of the tetrahedron.

This enables verification of the Theorem 2 by the method of descriptive geometry in connection with a dynamic geometry software.

ACKNOWLEDGMENTS

The research is partially supported by the grant of the University of South Bohemia GAJU 017/2013/S.

REFERENCES

- [1] A. Capani, G. Niesi and L. Robbiano. *CoCoA, a System for Doing Computations in Commutative Algebra*. <http://cocoa.dima.unige.it>
- [2] S. C. Chou. *Mechanical Geometry Theorem Proving*. D. Reidel Publishing Company, Dordrecht, 1987.
- [3] D. Cox, J. Little and D. O'Shea. *Ideals, Varieties, and Algorithms*. Second Edition, Springer, New York, Berlin, Heidelberg, 1997.
- [4] H. S. M. Coxeter, S. L. Greitzer. *Geometry revisited*. Toronto, New York, 1967.
- [5] GeoGebra [online]. <http://www.geogebra.at>
- [6] O. Giering. Affine and Projective Generalization of Wallace Lines. *J. Geom. and Graphics*, 1: 119–133, 1997.
- [7] G. M. Gruel, G. Pfister, H. Schönemann. *Singular 2.0. A Computer Algebra Sys-*

tem for Polynomial Computations. Univ. Kaiserslautern, 2001.

- [8] M. Guzmán. An Extension of the Wallace–Simson Theorem: Projecting in Arbitrary Directions. *Amer. Math. Monthly*, 106: 574–580, 1999.
- [9] B. Hunt. *The Geometry of Some Special Arithmetic Quotients*. Springer, New York, 1996.
- [10] E. Roanes–Lozano, E. Roanez–Macías. Automatic Determination of Geometric Loci. 3D-Extension of Simson–Steiner Theorem. In *AISC 2000*, (Campbell J. A., Roanes–Lozano, E. eds), Lecture Notes in Artificial Intelligence 1930, pages 157–173. Springer, New York, Heidelberg, 2001.
- [11] P. Pech. *Selected topics in geometry with classical vs. computer proving*. World Scientific Publishing, New Jersey, London, Singapore, 2007.
- [12] P. Pech. On Simson–Wallace Theorem and Its Generalizations. *Journal for Geometry and Graphics*, 9: 141–156, 2005.
- [13] T. Recio, H. Sterk and M. P. Vélez. Project 1. Automatic Geometry Theorem Proving. In: *Some Tapas of Computer Algebra* (Cohen, A., Cuipers, H., Sterk, H., eds.), Algorithms and Computations in Mathematics, Vol. 4, pages 276–296. Springer, New York, Heidelberg 1998.
- [14] D. Wang. *Elimination Practice. Software Tools and Applications*. Imperial College Press, London, 2004.
- [15] D. Wang. *Elimination Methods*. Springer-Verlag, Wien, New York, 2001.

ABOUT THE AUTHOR

Pavel Pech is working at the Dept. of Mathematics, Faculty of Education, University of South Bohemia in České Budějovice, Czech Republic. He is interested in geometry, proving theorems by computer and educational issues in mathematics.

AGGREGATION OF POLYHEDRON MODULES ON A FREEFORM SURFACE

Pierre CUTELLIC¹, Ursula FRICK², Thomas GRABNER², Rupert MALECZEK³

¹*Ecole Nationale Supérieure d'Architecture Paris-Malaquais, France*

²*[uto] Innsbruck, Austria*

³*University of Innsbruck, Austria*

ABSTRACT: This paper presents the ongoing research to obtain an aggregated structure consisting of trihedral polyhedral modules that approximate a given ideal design surface. The aim of this research is to develop geometric and computational design methods to generate such an assembly in order to erect spatial figurations.

In this paper a naming and construction method for the elements of the modules will be proposed. A categorization into regular and irregular assemblies will also be proposed. This paper can be seen as the beginning of a more detailed ongoing investigation.

Keywords: architectural geometry, approximation, faceting, modular system, polyhedral object, reciprocal systems.

1. GEOMETRIC CONSTRAINTS

The initial module of study is a non-trihedral "Normalon". "Normalons" are defined as those polyhedrons in which all edges are aligned with one of three mutually-perpendicular axes[1]. The Vertex N_0C (Figure 1) is the non-trihedral vertice.

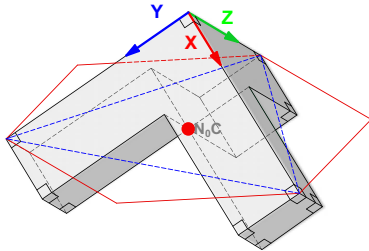


Figure 1: the initial module

In order to describe observed variations and preserved properties of the geometrical object, we will define a naming standard related to the geometrical properties and construction of such a Normalon inscribed within a tetrahedron (Figure 2), as well as to its topology.

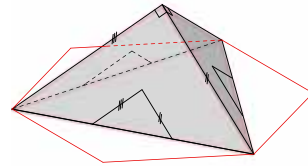


Figure 2: a regular tetrahedron as base for the initial module

1.1 Topology

As three axes meet in a single point and indicate the orientation of the module, we will name 4 main regions of the polyhedron as follows : N_X , N_Y , N_Z and N_0 (Figure 3).

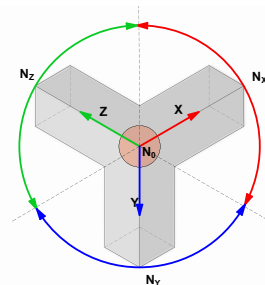


Figure 3: the regions defined by its axis

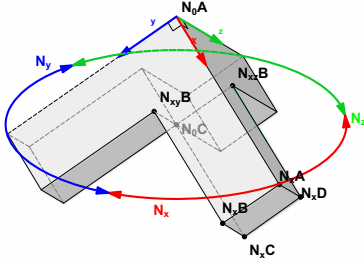


Figure 4: Vertices of one region (N_x) that form a “finger”

N_0 contains two vertices : N_0A corresponding to the top vertex of the inscribing tetrahedron, and N_0C corresponding to a vertex resulting from the truncation of the same tetrahedron which will be described hereafter. Each of the three other regions contain eight vertices among which two are mutually shared by the regions.

By rotating clockwise around each axis and starting from the vertex lying on the edge of the inscribing tetrahedron which also contains N_0A , we obtain N_{xA} , N_{xB} , N_{xC} , N_{xD} , N_{yA} , N_{yB} , N_{yC} , N_{yD} , N_{zA} , N_{zB} , N_{zC} and N_{zD} . We are then left with N_{x0A} , N_{y0A} , and N_{z0A} . The last three mutually shared vertices are called N_{xyB} , N_{yzB} and N_{zxB} . Edges will be named after their vertex boundaries and faces after their construction plane (Figure 4).

For their particular form and their function within the entire assembly we name the polyhedra's group of facets formed by these vertices fingers.

1.2 Construction

As previously mentioned, the normalon module is inscribed within a tetrahedron which will be later truncated by a self-similar module to give shape to another polyhedron. We name the four planes of the tetrahedron: T_{XY} , T_{YZ} , T_{ZX} , T_{00} (Figure 5).

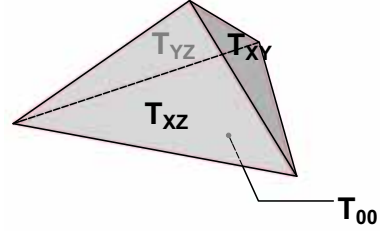


Figure 5: the tetrahedron's planes naming

In order to perform a two-dimensional compact tiling, a dual grid (hexagonal, triangular) is used to position each instance of the tetrahedron and results in the aggregation of one tetrahedron (T_0) with six others (T_1 , T_2 , T_3 , T_4 , T_5 , T_6) of a primary surrounding ring.

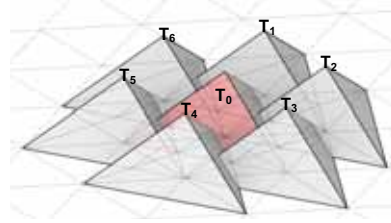


Figure 5: Tetrahedron with surrounding Tetrahedrons

In order to obtain all faces of the polyhedron, we need to use the intersection of 3 neighbors of the first crown on the faces T_{XY} , T_{YZ} , and T_{ZX} . The three remaining faces are then obtained by using six neighbors of the second ring. This fully defines the truncation of one polyhedron which all faces are consequently planar and we name a normalon and the nine necessary tetrahedras for the truncation (T_1 , T_2 , T_3 , T_4 , T_5 , T_6 , T_7 , T_8 , T_9 , T_{10} , T_{11} , T_{12}).

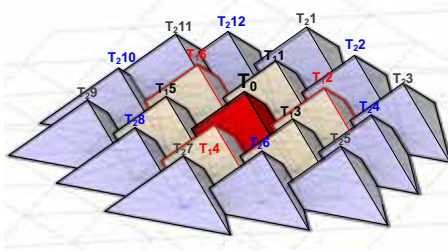


Figure 6: Tetrahedron the two surrounding rings.

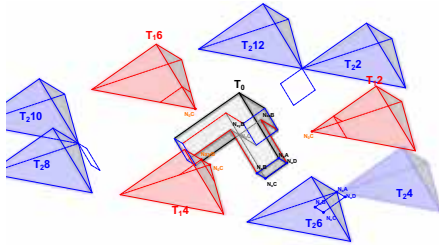


Figure 7: Tetrahedron with the nine tetrahedras of truncations of the two surrounding rings.

The relation between the surrounding tetrahedrons and faces can be seen in Figure 7. While the Tetrahedrons T_{12} , T_{14} , T_{16} , T_{22} , T_{24} , T_{26} , T_{28} , T_{210} and T_{212} in blue are truncating the caps of the fingers, are T_{12} , T_{14} and T_{16} colored red, responsible for the inner edges of the finger. The bottom edge of the fingers is therefore the result of the intersection of the tetrahedrons T_{12} , T_{14} and T_{16} .

Each face will then be named after the plane on which it lies, whether or not the plane belongs to another tetrahedron. We will have twelve faces in total (Figure 8)

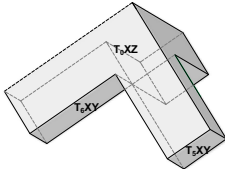


Figure 8: Faces named after their relation to the neighbors

1.3 Properties

This is the topological notation we obtain : $25v-24e-12f$ (25 vertices, 24 edges, 12 faces). The regular normalon resulting from the compact tiling of self similar regular tetrahedrons. The ideal regular element consists of planar faces and has only 90 degree angles. The entire volume can be described by seven identical cubes(Figure 9). Consequently, edges of the caps, fingers and heaps are equal or proportional to the module of the cube.

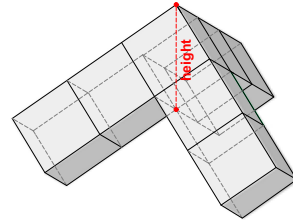


Figure 9: the ideal Module consisting from seven identical Cubes

A very important factor of the element is its height. In the ideal Module the height is the diagonal of the cube and is oriented normal to the plane of the dual grid. The height and its direction is the Vector between the Points N_0C and N_0A and is named main Orientation.

1.4 Aggregation

The generation of the tetrahedrons and the truncation method, in combination with the proportions are very important for the assemblage of the modules to a structure. In our case, the modules can be assembled by sliding the fingers of the modules into each other in a particular order to form an assembly. As shown in Figure 10, the upper faces of each fingers are touching with the lower third the upper half of the inner faces of the neighboring finger. The Modules can only be assembled with this sliding strategy when the section of the fingers are smaller or equal to the one of their heap (Figure 11).

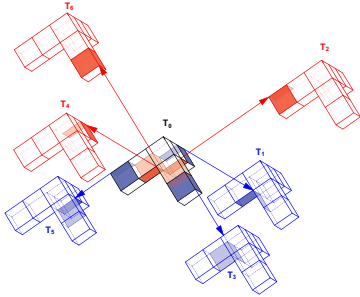


Figure 10: The touching faces of the modules and their sliding direction.

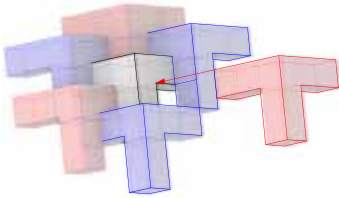


Figure 11: a more detailed view of the assembly strategy

These uniform modules, aligned perpendicular to one another, can only approximate planar uniform surfaces. In aggregation, the polyhedral objects represent a faceted volumetric planar body (Figure 12).

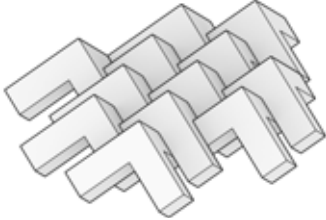


Figure 12: Modules assembled to a flat surface

In addition, these assemblies, due to their configuration, and with a set boundary condition, work as self-interlocking system. This system was used during the AA-Summer School located in Ivrea [3], in 2010 to build a pavilion (Figure 13) from four plane surfaces.



Figure 13: the pavilion named ziggurat, built in Ivrea

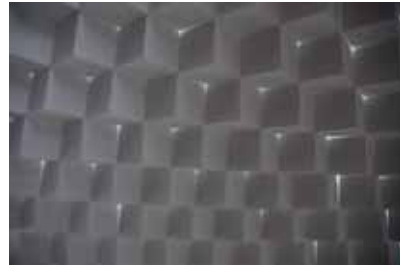


Figure 14: the faceted interior

2. SURFACE APPROXIMATIONS

In order to apply the module on a double-curved design surface, the geometric constraints associated with the alignment of the axes/edges are to be abdicated. Using the same geometrical typology, but applying a module with a degree of freedom in the edge alignment, the approximation of freeform surfaces with a trihedral polyhedral aggregation is achievable.

2.1 Approximation strategies

For a better explanation, we consider the assembly of the ideal Module under a different constrain. If we wish to approximate a flat surface with the presented modules then we have to generate the tetrahedrons and its underlying dual grid in relation to a surface. We propose to generate a grid from intersecting equal sided hexagons, representing the bases N_0 of the tetrahedrons, on this surface and use then the surface normal or another chosen vector as Main Orientation to generate the tips of the tetrahedrons with a favored height.

In the further step the tetrahedrons are truncated considering the constraints explained above.

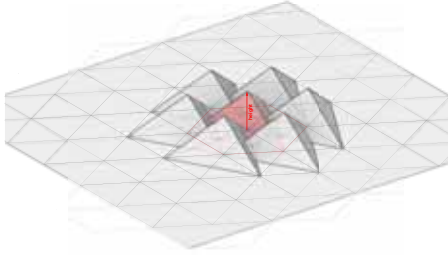


Figure 15: a flat surface approximated with the tetrahedrons

For curved surfaces, we propose a similar procedure. To address the packing and fitting accuracy of the original planar system, each module's shape has to be correlated to the six neighboring cells of the first crown and three of the second crown and the local surface curvatures. Moreover, there is a strong association between the surface curvature, the module size, the module edge alignment, the modules main orientation, and the structural capacity of the assembly. These assets and conditions allow for a variation in the system to approximate the surface within a certain threshold.

It is an explicit aim of the entire research to investigate and understand these constraints. At the actual state of research, the grid is generated on curved surfaces, with the UV information of the surface.

2.2 Regular Assemblies

The normalon is a polyhedral figure, that is per definition a regular module. In the context presented here we define regular elements, when all three fingers have an identical form. We name spatial figurations from identical regular modules regular assemblies.

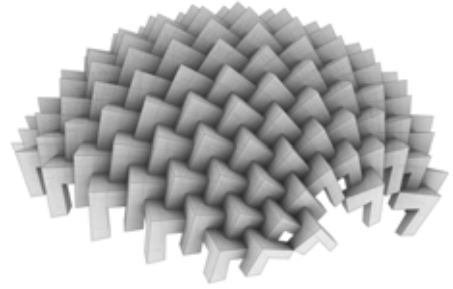


Figure 16: a regular assembly on a double curved surface

As described above, applying the modules to curved surfaces generate deformed modules, that are constrained with its neighbors. The deformation is influencing mainly the form of the fingers. When using the surface normals as main orientation, constant gaussian curvature generates identical finger forms.

Even if identical, the height is influencing the entire topology of the fingers and therefore the modules themselves (Figure 17). This can be an important information for designers, that will work with this system for architectural purpose. Regular assemblies generated from regular modules, can therefore only approximate surfaces with a constant gaussian curvature.

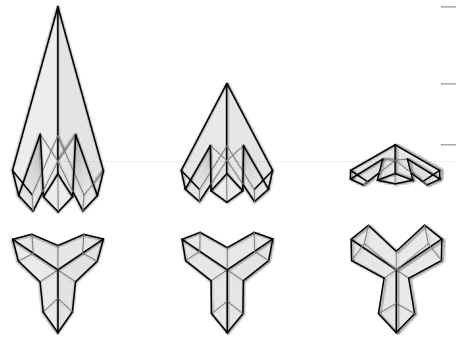


Figure 17: three modules from constant gaussian curvature with different heights

2.3 Irregular Assemblies

One of the main intent of this research is the approximation of curved surfaces for different purposes. As most curved surfaces, do not have a constant gaussian curvature, the proposed approximation strategies will generate assemblies form irregular modules, with different finger forms. We name this assemblies irregular assemblies.

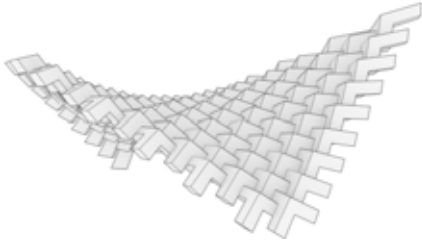


Figure 18: an irregular assembly approximating a hyperbolic paraboloid surface

The generation of an aggregation with all different modules is relatively easy, but the production of these modules can be very difficult and cost intensive. One solution for designers could be the use of extruded surfaces along an axis, in order to minimize the number of different modules within an assembly, as the Elements are identical in extrusion direction. If a planar arc is extruded and approximated, the entire assembly will consist of identical elements, with two mirrored fingers in the arc's plane and one individual finger in the extrusion's direction (Figure 19).

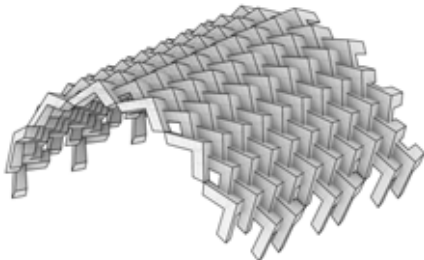


Figure 19: an extruded arc approximated with identical modules

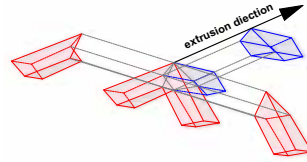


Figure 20: the used module from Figure19, the blue Finger is in extrusion direction

As most assemblies will be irregular assemblies, one future task of this research is to define strategies and techniques that will allow as many identical elements as possible on a given design surface.

3. CONCLUSION

We presented a possible method for the aggregation of polyhedron modules on a freeform surface into spatial figurations. We described a system, and showed the relations of the single modules to each other. We proposed a technique to approximate surfaces and made also a proposition to define regular and irregular assemblies for a later and more intense research.

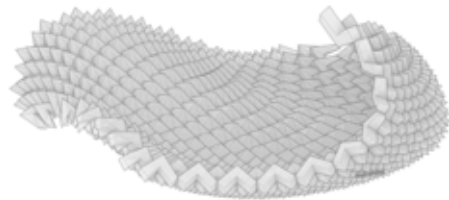


Figure 21: a double curved surface with an irregular aggregation

4. OUTLOOK

The ongoing research focuses on the relationship between these geometrical properties and the integrity of the structure, as well as on the development of a digital automation process for the geometry modifications. The stability and self-interlocking capacity of the system will be further discussed and documented through physical models and prototypes.

4.1 Self interlocking Systems

Self-interlocking systems are very rare in architecture and design, even if known since a long period. But since a few years, new technologies and approaches allowed some interesting developments within this field [1,2,3]. The aim for this particular aggregation system is to find a set of possible boundary conditions, rules and geometric constraints, that allow the generation of these types of structures as self interlocking assemblies. Ongoing investigations are made with physical simulation software (Figure 22).

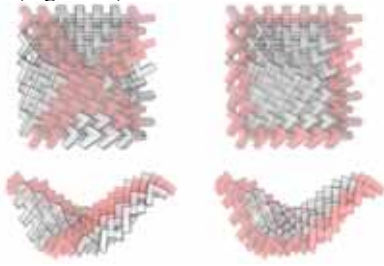


Figure 22: two test configurations for the investigation of self interlocking – the fixed elements are marked red

4.2 Fabrication

The presented module is a very complex polyhedron that needs a further investigation on its fabrication possibilities. The built example in Ivrea was fabricated from extruded polystyrene foam. As the used module was a normalon it could be produced from rectangular elements.

As we have different polyhedron geometries on approximated freeform surfaces, the fabrication will be more delicate. This investigation will be focused on possible materials, digital fabrication methods, and assembly strategies. The assembly strategy will be a key element for this part of the future research. The freely formed fingers will probably block each other during the assembly of several modules into a structure. One solution could be the fabrication of the modules “on site” or to search for an optimization of the geometrical constraints, so the different elements can be prefabricated and assembled on site.

ACKNOWLEDGMENTS

This research became possible through the financing help of the France-Focus from the University of Innsbruck. At the workshops held in Paris and Innsbruck, some highly motivated students as Clemens Jenny, Alexander Topf, Sylvain Usai, Jim Rhone were involved and giving a helping hand for this investigation.

We want also to mention the AA-Summer School in Ivrea 2011 and Tomaso Franzolini. This event allowed us to build the initial prototype for this particular research.

REFERENCES

- [1] P. Varrley. Resolution Sequences for Geometric Beautification. Proceedings of the World Congress on Engineering 2010 Vol. I, WCE 2010, London, U.K.
- [2] M. Rippmann and P. Block. New Design and Fabrication Methods for Freeform Stone Vaults Based on Ruled Surfaces. In *Gengnagel, Kilian, Palz, Scheurer (eds), Computational Design Modelling*, pages 181–189. Springer, Berlin, 2011.
- [3] M. Brocato and L. Mondarini. Geometric methods and computational mechanics for the design of stone domes based on Abeille’s bond. Ceccato, Hesselgren, Pauly, Pottmann, Wallner (eds). *Advances in Architectural Geometry 2010*, pages 149-162. Springer, Wien, 2010.
- [4] O. Tessmann. Interlocking Manifold – Kinematically Constrained Multi-material Systems. Hesselgren, Sharma, Wallner, Baldassini, Bompas, Raynaud(eds). *Advances in Architectural Geometry 2012*, pages 269-278. Springer, Wien, 2012.
- [5] AA-SummerSchool entitled “Factory Futures”, Ivrea, Italy. 2010. Director Tomaso, Franzolini, co-organised by Pierre Cutelic and Rupert Maleczek.

ABOUT THE AUTHORS

1. Pierre Cutelic is founder and principal of The Computational Monkeys, Adjunct-Assistant Professor at ENSA Paris-Malaquais and Scientific Collaborator at EPFL CNPA.

2. Ursula Frick is member and cofounder of the design and research collective [uto]. Her research focuses, along with the generation and analysis of complex geometries, on the development of formal strategies and data driven design approaches.

3. Thomas Grabner is member and co-founder of the design and research collective [uto]. He is an architect and specialist in computational design with several years of practice and teaching experience. His research focuses, along with the generation and analysis of complex geometries, on the development of formal strategies, data driven design and implementing and evaluating the physical matter of a design.

4. Rupert Maleczek, is researcher and architect that has teaching and research experience in folded systems, deployable structures, digital fabrication and geometry. He is member and co-founder of the research and design collective Archiwaste, and Founder and principal of Archispass.org.

THE AID OF GEOGEBRA TO TEACH AND LEARN DESCRIPTIVE GEOMETRY AT THE FACULTY OF OPERATION AND ECONOMICS OF TRANSPORT AND COMMUNICATIONS AT THE UNIVERSITY OF ŽILINA IN ŽILINA, SLOVAKIA

Viera ČMELKOVÁ

The University of Žilina in Žilina, Slovakia

ABSTRACT: The University of Žilina in Žilina was established in 1953 by separation from the Czech Technical University in Prague. It has undergone many transformations during its history. In 1996 it was, eventually, renamed from the University of Transport and Communication to the University of Žilina in Žilina.

The crux of the study and scientific research of the Faculty of Operation and Economics of Transport and Communications is economics and operation of various kinds of transport (road transport, rail transport, water transport and air transport), postal technologies and services, forwarding and logistics, economics and management of enterprise, electronic business and management.

Most students who are starting their studies at the Faculty of Operation and Economics of Transport and Communications come from secondary schools with focus on economics. However, the primary target of the economic studies on the Faculty of Operation and Economics of Transport and Communications lies in technical fields, which places requirements on students to obtain basic skills and knowledge about object relationships in three-dimensional Euclidean space and methods of projection of three-dimensional Euclidean space to the plane (Monge method, orthogonal axonometry).

After several major changes that the Slovak school system has undergone over the last few years (secondary-school leaving examination in mathematics is no longer required, each school has its own program of study, ...), the asymmetry in mathematical skills and knowledge, and in geometrical skills and knowledge especially, has increased significantly. This has a very disruptive effect to the teaching of the first year undergraduate courses on related subjects. Consequently, the use of dynamic geometry software like GeoGebra is highly desirable.

Rather than solving the problems of planar constructions of descriptive geometry, the students use GeoGebra for understanding the topics by analysing the applets prepared by us, the lecturers. They can do so not only during the lectures, but also anytime later, dedicating as much time to the subject as they require. We, as teachers, see the benefit of GeoGebra in several aspects. It is a very user-friendly software; for example, there are options available to set break points in construction protocol, or to set some fractional constructions invisible, so there is not too much objects in the structure and the construction is much more transparent and comprehensive. One of the biggest benefits is the possibility to share GeoGebra applets as HTML5 applets on GeoGebraTube, so students do not need to install GeoGebra on their PC. One of the latest benefits is GeoGebra application for Android, so students could study everywhere. Last but not least, GeoGebra is a free software, so students do not need to pay for it.

While our utilisation of Geogebra does not aim at becoming proficient in descriptive geometry, it is useful for both teaching and learning of some of its introductory problems and constructions used in methods of projection of three-dimensional Euclidean space into plane.

Keywords: GeoGebra, Descriptive Geometry, Orthogonal Axonometry, Monge Method, Geometry

1. THE UNIVERSITY OF ŽILINA IN ŽILINA

The University of Žilina in Žilina was established in 1953 by separation from the Czech Technical University in Prague under the name Railway College. In 1959 it was renamed to the University of Transport and in September 1960 was relocated from Prague to Žilina. It has three faculties then:

- Faculty of Operation and Economics of Transport
- Faculty of Mechanical and Electrical Engineering
- Military Faculty

University has undergone many transformations during its history and in 1996 was constituted as the University of Žilina in Žilina. Currently the university has seven faculties:

- Faculty of Operation and Economics of Transport and Communications
- Faculty of Mechanical Engineering
- Faculty of Electrical Engineering
- Faculty of Civil Engineering
- Faculty of Management Science and Informatics
- Faculty of Special Engineering
- Faculty of Humanities



Figure 1: Our students reached Slovak record “The most students photography” [1]

University also includes a number of departments and institutes (CETRA - Centre of Transport Research, Institute of Forensic Engineering, Institute of Competitiveness and Innovations, Institute of High Mountain Biology, Institute Research Centre, Institute of Information and Communication Technologies and other).



Figure 2: University of Žilina in Žilina. [5]

The university provides education at all three levels of higher education (Bachelor's degree, Engineer/Master's degree and Doctoral degree) in full time and part-time forms. With its scientific and foreign activities based on cooperation with domestic and foreign companies and institutions, with almost 130 accredited educational programmes and with number about 11,000 students studying in all forms of study takes University of Žilina in Žilina a prominent place in the educational area in Slovakia.

2. THE FACULTY OF OPERATION AND ECONOMICS OF TRANSPORT AND COMMUNICATIONS

Faculty of Operation and Economics of Transport and Communications is the part

of University of Žilina in Žilina since the beginning. The original name was the Faculty of Operation and Economics of Transport. In 1976 the faculty received the current name.



Figure 3: Railway laboratory. [4]

At present, the Faculty consists of seven departments:

- Department of Road and Urban Transport
- Department of Economics
- Department of Quantitative Methods and Economic Informatics
- Department of Air Transport
- Department of Communications
- Department of Water Transport
- Department of Railway Transport

The faculty provides education at all three levels of higher education (Bachelor's, Engineer's and Doctoral degree) in full-time and part-time form of study in 22 accredited educational programmes for more than 2,670 students. As it is evident from the name of the faculty, science, research and education target lies in field of operation and economics of transport and communications.

3. GEOMETRY

The subject Geometry is being taught at the Faculty of Operation and Economics of Transport and Communications in the educational programs

- 3772 7 14 air transport
- 3772 7 15 road transport

- 3772 7 16 railway transport
- 3772 7 17 water transport
- 3772 7 19 professional pilot

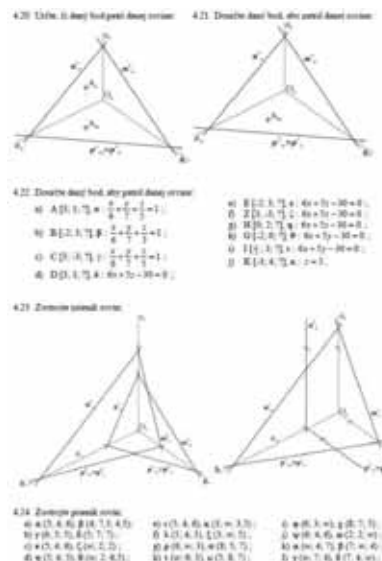


Figure 4a: Workbook.

The subject geometry is included in the 1st semester of the first year of Bachelor's degree of study. The extent of the study is 2 hours of lectures and 2 hours of practice per week (the length of semester is 13 weeks). The form of study is full-time study.

The aim of the subject Geometry is to introduce students to the basics of descriptive, analytic and differential geometry, and in liaison with the mathematics to develop a link between the analytic expression and graphic representation of objects, as well as to support the development of space visualization and graphic expression.

The topics of the study are:

- Euclidean plane, Euclidean space
- The coordinate system in Euclidean plane and Euclidean space
- Conic - mainly ellipse

- Linear transformation
- The axial affinity – the view of a point, line, polygon, ellipse
- The axial collineation – the view of a point, line, polygon
- The projection – central projection, parallel projection (basic characteristics)
- The orthographic projection to the one plane – the plan view of a point, line, finding the real measure of the segment
- Monge method – the top and front view of a point, a line, a traces of a plane, a simple body with the base lying in horizontal or frontal plane
- Orthogonal axonometry – the axonometric view of a point, a line, a traces of a plane, a polygon and a circle lying in a ground plane, basic tasks of position of elements, the axonometric view of simple bodies with bases lying in ground plane, also plane sections of this bodies (except cone) and the intersection of the line with this bodies
- Analytic representation of a planar curve, the tangent and asymptote of the planar curve.

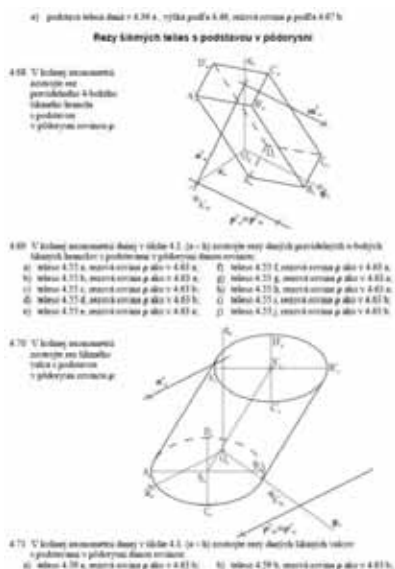


Figure 4b: Workbook.

4. GEOGEBRA APPLETS

GeoGebra is a freely distributable general-purpose mathematical software that can be utilized in many fields of mathematical education and model creation processes. It is currently available in more than 50 languages and offers a very user-friendly interface. This program was previously meant to be used for Euclidean geometry problem-solving, e.g. to draw a point, a segment, a midpoint of a segment, an axis of a segment, an intersection of two objects, a ray, a line, a line perpendicular to another line passing through a given point, a line parallel to another line passing through a given point, construction of an angle of a given size, angle axis construction, work with vectors, construction of a regular polygon, circle and circular curves given in several ways, creation of a conic given by its defining elements, i.e. by five points, construction of a polar line, displaying objects in identical transformations (rotation, displacement, axial symmetry, ...) in dilatations or in circular inversion, measuring object areas and distances. GeoGebra later expanded into other fields of mathematics, including mathematical analysis, algebra, statistics, table calculations (e.g. examining a graph of a function, probability calculations, linear regression etc.).

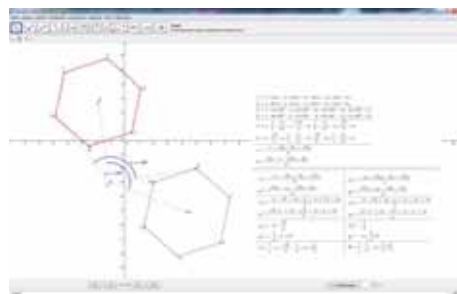


Figure 5: Image of a hexagon rotated around the common point H. [12]

All geometrical objects rendered in GeoGebra may be expressed by means of algebraic

formulae, since it is possible to insert a coordinate system into the drawing plane, and then possible to edit the algebraic expression later according to the users' needs. It is possible to move the objects, which affects the formula accordingly. The program also allows enriching the drawing plane with scrollbars, buttons, checkboxes and text areas, allowing the student himself to modify the objects displayed on the drawing plane and leading to student's better understanding of the given topic.

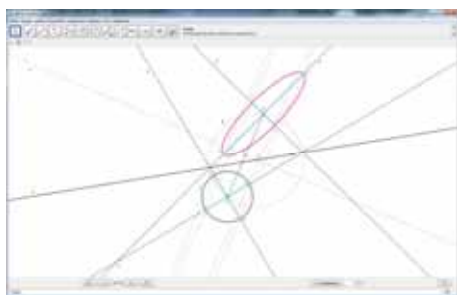


Figure 6: Affine image of a circle. [8]



Figure 7: Affine image of an ellipse. [7]

The authors of applets can also use the option to create an animation. Another useful feature is the ability to show only some of the objects (thus making the construction much more comprehensible). Intermediate GeoGebra users may appreciate the ability to create their own procedures, which in turn simplify their work in creating applets (e.g. the Rytz method procedure). Advanced users and apt program-

mers will certainly laud the option to code their own elements since GeoGebra is an open-source software available freely for non-commercial purposes. Very useful is possibility to make mathematical forms and text using computer programming language TeX without necessity of installing TeX. Another advantage of GeoGebra which should not be forgotten about is the option to export the constructions and applets into HTML format so that the students are not forced to install GeoGebra when only viewing their study materials created in the programme. These HTML applets are uploaded to the GeoGebra Tube website and are freely available to anyone. Beginner users can acquaint themselves with the program in its installation-free online version GeoGebra Start.



Figure 8: Ground view and front view of an oblique cone with the base circle lying in the ground plane. [11]

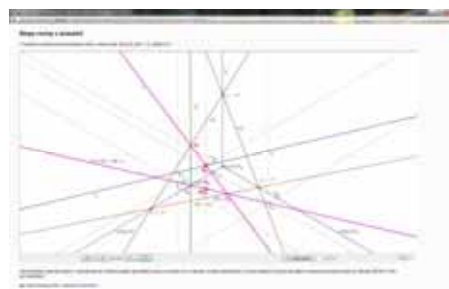


Figure 9: Isometric view of point of intersection of line and plane. [9]

In December 2013, the authors of GeoGebra launched versions for tablets and iPhones in addition to the desktop edition of the software, which allows students to study materials created by teachers anywhere, anytime.

A version of GeoGebra with three-dimensional objects mapping and three-dimensional object transformations has recently gone into development. It is currently possible to produce an axonometric or perspective view of an object, along with producing its plan view, front view and side view. The beta version features the following 3D transformations: symmetry about a plane, line, point, rotation about a line, translation in the direction of a given displacement vector and dilation (enlargement) according to a given point.



Figure 10: Orthogonal axonometric view of plane section of an oblique cylinder. [10]

What we, the teachers of descriptive geometry, appreciate about GeoGebra is its dynamic nature. The possibility of study construction applet line after line, the option of having auxiliary lines hidden from the eye of the student, the opportunity whenever pause the dynamic applets, move frame-by-frame back and forward as needed.

We have created applets for the students of our faculty to help them understand the principles of identical and non-identical transformations along with the principles of projecting three-dimensional Euclidean space into a plane.

Applets are also available to revise identical and non-identical transformations of a plane (rotation, axial symmetry, etc.), applets to construct conics, to create images of objects in axial affinity and central collineation, applets to construct object images using Monge method and using orthogonal axonometry method, plane sections of objects, points of intersection of lines and simple objects mapped via orthogonal axonometry method.

5. CONCLUSIONS

While our utilisation of Geogebra does not aim at becoming proficient in descriptive geometry, it is useful for both teaching and learning of some of its introductory problems and constructions used in methods of projection of three-dimensional Euclidean space into plane.

ACKNOWLEDGMENTS

This paper was designed in response to the project co-funded by the EU called "The quality of education and human resources development as pillars of a knowledge society at the Faculty PEDAS, ITMS 26110230083".

REFERENCES

- [1] čas.sk, <http://www.cas.sk/galeria/310300/no-vy-rekord-zmerali-zrnkami-ryze-az-1-787-s-tudentov-na-jednej-fotke?foto=0> [June 10th, 2014]
- [2] Faculty of Operation and Economics of Transport and Communications, <http://fpedas.uniza.sk/sk/> [June 10th, 2014]
- [3] Geometria – source of information, <https://vzdelavanie.uniza.sk/vzdelavanie/plainfo.php?kod=162737> [June 10th, 2014]
- [4] Katedra železničnej dopravy, <http://kzd.uniza.sk/index.php/foto/category/4-sucasny-labak> [June 10th, 2014]
- [5] mojaBystrica.sk, http://www.mojabystrica.sk/subportals2/elektrotech-nicka-fakulta-zilinska-univerzita-v-ziline-banska-bystrica.phtml?id_temy=808&action=

- [clanky&id_clanku=111779&p_sekcia=](#)
[June 10th, 2014]
- [6] University of Žilina in Žilina,
<https://www.uniza.sk/menu/inc.php?id=130>
[June 10th, 2014]
- [7] https://vzdelavanie.uniza.sk/moodle/file.php/425/Geogebra/html/Afinita/afinita_elipsa_1.html [June 10th, 2014]
- [8] https://vzdelavanie.uniza.sk/moodle/file.php/425/Geogebra/html/Afinita/afinita_kruznic_2.html [June 10th, 2014]
- [9] https://vzdelavanie.uniza.sk/moodle/file.php/425/Geogebra/html/Axonometria/prienik_priamky_s_rovinou_izometria.html
- [10] https://vzdelavanie.uniza.sk/moodle/file.php/425/Geogebra/html/Axonometria/rez_vallec_v_pi_01_sikmy.html [June 10th, 2014]
- [11] https://vzdelavanie.uniza.sk/moodle/file.php/425/Geogebra/html/Mongeovka/sikmy_kuzel_v_pi.html [June 10th, 2014]
- [12] https://vzdelavanie.uniza.sk/moodle/file.php/425/Geogebra/html/Transformacie_v_E2/otocenie_sestuholnik_okolo_H.html [June 10th, 2014]

ABOUT THE AUTHORS

1. Viera Čmelková works as an assistant professor at the Department of Quantitative Methods and Economic Informatics (The Faculty of Operation and Economics of Transport and Communications, University of Žilina in Žilina, Slovakia). She teaches Geometry and Mathematics. In scientific research, her foci include teaching methodology in mathematics (esp. geometry), geometry and topology and partly also mathematical statistics and its application.

THE ANALOGUE OF THEOREMS RELATED TO WALLACE-SIMSON'S LINE IN QUASI-HYPERBOLIC PLANE

Ana SLIEPČEVIĆ¹ and Ivana BOŽIĆ²

¹University of Zagreb, Croatia

²Polytechnic of Zagreb, Croatia

ABSTRACT: The quasi-hyperbolic plane is one of nine projective-metric planes where the absolute figure is the ordered triple j_1, j_2, F . It is dual to the pseudo-Euclidean plane. It is known for the fact that a pencil of parabolas, in the Euclidean and pseudo-Euclidean plane, can be set according to lines a, b, c . The focus points of all parabolas in the pencil lie on the circle circumscribed to the triangle given by lines a, b, c . The connection between the pencil of parabolas, Wallace-Simson lines and Steiner deltoid curve are studied and proved in [2]. Analogues theorems are valid in the pseudo-Euclidean plane. In this paper the dual theorems will be proved in quasi-hyperbolic plane.

Keywords: Quasi-hyperbolic plane, pencil of parabolas, Wallace-Simson line, point A

1. INTRODUCTION

It is known that there exist nine geometries in plane with projective metric on a line and on a pencil of lines which are denoted as Cayley-Klein projective metrics. Hence, these plane geometries differ according to the type of the measure of distance between points and measure of angles which can be parabolic, hyperbolic, or elliptic. The plane geometry with hyperbolic measure of distance and parabolic measure of angle is denote as the quasi-hyperbolic plane (further in text qh-plane). Furthermore, each of the Cayley-Klein projective metrics can be embedded in the real projective plane $\mathcal{P}_2(\mathbb{R})$ where then the metric is induced by an absolute figure which is given as non-degenerated or degenerated conic [5], [6], [7]. The absolute figure in the qh-plane is a real point F and a pair of real lines j_1 and j_2 incidental with F . In order to simplify the constructions the extended model of qh-plane where all points and lines of the qh-plane embedded in the real projective plane $\mathcal{P}_2(\mathbb{R})$ are observed. In [1] some basic geometric notions of the qh-plane are introduced, also the classification of qh-conics with respect to their position to the absolute figure is given and some basic con-

structions are presented. In [2] we have studied and proved the connection between the pencil of parabolas, Wallace-Simson lines and Steiner deltoid curve. It is not difficult to conclude that in the pseudo-Euclidean plane the analogous theorems are valid. In this paper we will not proved them. The main aim of the paper is to proof the dual of above mentioned theorems in the qh-plane, by using the notions defined in [1]. Following notions need to be highlighted :

- The lines incidental with the absolute point F are called **isotropic lines**.
- Two points A and B in qh-plane are called **perpendicular points** if they lie on a pair of isotropic lines a and b that are in harmonic relation with the absolute lines j_1 and j_2 .
- A **qh-circle** is a qh-conic for which the tangents from the absolute point F are the absolute lines j_1 and j_2 .
- A **qh-parabola** is a qh-conic passing through the absolute point F i.e. both isotropic tangent lines coincide. A **special**

parabola is a qh-parabola whose isotropic tangent is an absolute line.

- The **directrices** of a qh-conic are (non-absolute) lines incident with the isotropic points of the qh-conic, i.e. lines incidental with the intersection points of the qh-conic with the absolute lines j_1 and j_2 . A qh-conic can have none, one, two or four directrices f_i , $i \in \{1, 2, 3, 4\}$. The directrices are the dual of the pseudo-Euclidean focuses.

2. THE ANALOGUE OF THEOREMS RELATED TO WALLACE-SIMSON'S LINE

Theorem 1. *Let M , N and P be three non-isotropic points, incidental with three different isotropic lines. If the pencil of qh-parabolas is set according to points M , N and P then the envelope of the directrices of the qh-parabolas is the qh-circle inscribed in the trilateral mnp .*

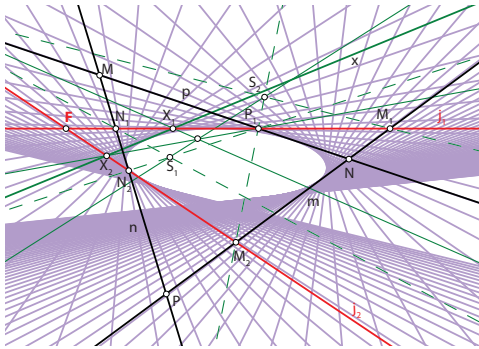


Figure 1:

Proof: Let the point X_1 be an arbitrary chosen point on a absolute line e.g. $X_1 \in j_1$. The qh-parabola from the given pencil is set according to points F , M , N , P and X_1 . Let the intersection point of the qh-parabola and the absolute line j_2 be denoted as X_2 . The line $x := X_1X_2$ is its directrix (see Fig. 1). The ranges of isotropic points

(j_1) and (j_2) are determined by all the parabolas from the given pencil ($X_1 \in (j_1)$, $X_2 \in (j_2)$). The ranges (j_1) and (j_2) are projectively linked, and the result of its correspondence will be the 2nd class curve [8]. In the given pencil of qh-parabolas there are two special qh-parabolas in the cases when X_1 or X_2 coincides with F , respectively. Its directrix coincides with the absolute line that is different from its isotropic tangent line. Therefore, the absolute lines are the tangent lines of mentioned 2nd class curve i.e. the envelope of the directrices of the qh-parabolas from the pencil is a qh-circle. The sides MN , NP , MP of the given trilateral are the tangent lines to a qh-circle and they coincide with the directrices of three degenerated qh-parabolas from the pencil. \square

By using the results of the previous theorem we can prove the following dual of Wallace - Simsons theorem [2].

Theorem 2. *Let mnp be a trilateral with non-isotropic sides m , n , p and non isotropic vertices M , N , P and k its inscribed qh-circle. Let x be an arbitrary tangent line to k , and M_1 , N_1 and P_1 its respective perpendicular points to the vertices M , N and P . Then the lines M_1M , N_1N and P_1P intersect at the point A .*

Proof: Without loss of generality, to simplify the construction, qh-circle is represented with the Euclidean circle. Notice that the pencil of qh-parabolas can be determined by vertices M , N and P . By previous theorem it follows that the qh-circle, inscribed in trilateral mnp , is the envelope of the directrices of the qh-parabolas from the pencil. In the given pencil there are three degenerated qh-parabolas, a pair of lines (MN, PF) , (NP, MF) and (MP, NF) . Its directrices coincide with the lines MN , NP and MP , respectively. The tangent line x is the directrix of the qh-parabola, denoted as p_1 , that passes through the vertices M , N , P . The qh-parabola p_1 is incidental with the intersection points of the directrix x with the absolute lines j_1 and j_2 ,

denoted as X_1 and X_2 . The point F_1 is the focus of the qh-parabola (see Fig. 2). By using the definition of the pedal transformation in the qh-plane in [3] it is proved that the pedal qh-curve of a qh-parabola, with respect to an arbitrary polar line of the pedal transformation, is a 3^{rd} class curve. If the polar line of the pedal transformation coincide with the directrix x , the 3^{rd} class pedal curve degenerates into three points i.e. three pencils of line (X_1) , (X_2) and the pencil at the vertex of a qh-parabola. In our case point A is the vertex of the qh-parabola determined by points M , N , P and directrix x . \square

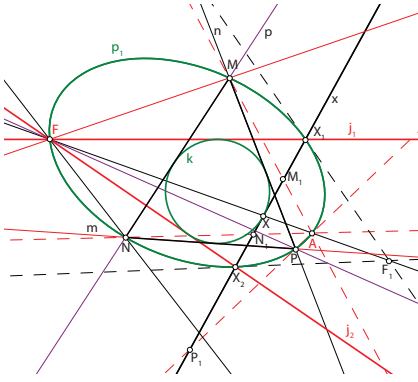


Figure 2:

Remark 1. *The curve in the qh-plane is an entirely circular if its isotropic tangent lines coincide with the absolute lines, and there are no other isotropic tangent lines.*

Theorem 3. *All the points A from theorem 2 lie on an entirely circular 4^{th} class cubic (see Fig. 3).*

Proof: Let mnp be a trilateral and k its inscribed qh-circle. According to construction of a point A , with respect to an arbitrary chosen tangent line x to k , the result of $(1-2)$ -correspondence between two pencils of lines (M) and (N) (or (M) and (P) or (N) and (P)) is a curve consisting of all points A . According to

$(1-2)$ -correspondence each line from the first pencil corresponds to two lines from the second pencil, conversely [4]. Hence, let m_1 be an arbitrary line from the pencil (M) , and the point M_1 a unique point on m_1 perpendicular to the point M (see Fig. 3). Let the tangent lines to k from the point M_1 , be denoted as t_1 and t_2 . Let the points $N_1 \in t_1$ and $N_2 \in t_2$ be perpendicular to the point N . The lines $n_1 := NN_1$ and $n_2 := NN_2$ are corresponded to the line $m_1 \in (M)$. Let the point of intersection of the lines n_1, n_2 with the line m_1 be denoted as A_1 and A_2 , respectively. The points A_1 and A_2 are correspond to the tangent lines t_1 and t_2 . The result of this correspondence between the pencils of lines (M) and (N) is a 4^{th} order curve. Since, the line MN is corresponding to itself, the 4^{th} order curve degenerates into a cubic and the line MN . Furthermore, if the line $m_1 \in (M)$ is incidental with the absolute point F , then the point M_1 coincides with F , the lines t_1 and t_2 coincide with the absolute lines j_1 and j_2 and the points N_1 and N_2 coincide with F . Therefore, the point F is a double point of cubic, so it is 4^{th} class cubic. Since, on the absolute lines j_1 and j_2 there are no other points A except an absolute point F , the absolute lines are the tangent lines to cubic at the point F . Hence, the cubic is entirely circular [3]. Notice that theorem 3 is the dual of the Steiners deltoid curve theorem [2] \square

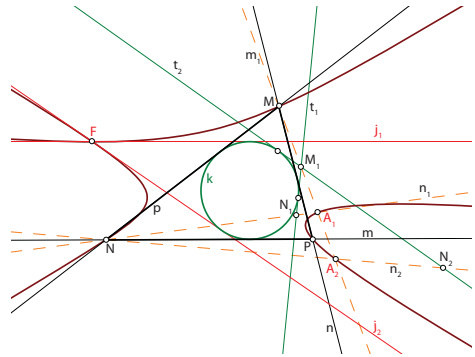


Figure 3:

Corollary 1. *The vertices of the qh -parabolas from the pencil lie on an entirely circular 4^{th} class cubic.*

3. CONCLUSIONS

The main aim of the paper is to proof the dual results of Wallace-Simpson line and Steiners deltoid curve theorem in the quasi-hyperbolic plane, by using the notions defined in [1].

REFERENCES

- [1] A. Sliepčević, I. Božić, and H. Halas. Introduction to the planimetry of quasi-hyperbolic plane. *KoG 17, Zagreb*, 58–64, 2013.
- [2] A. Sliepčević and I. Božić. Steiner curve in a pencil of parabolas. *KoG 16, Zagreb*, 13–15, 2012.
- [3] A. Sliepčević, H. Halas and I. Božić. Pedal curves of conics in quasi-hyperbolic plane. *Manuscript*.
- [4] H. Weileitner. Theorie der ebenen algebraischen Kurven hherer Ordnung. *G. J. Gschensche Verlagshandlung, Leipzig*, 1905.
- [5] I. M. Yaglom, B. A. Rozenfeld and E. U. Yasinskaya, Projective metrics. *Russ. Math. Surveys, Vol. 19, No. 5*, 51–113; 1964.
- [6] N. M. Makarova. On the projective metrics in plane. *Učenyje zap. Mos. Gos. Ped. in-ta (Russian)* 243, 274–290, 1965.
- [7] M. D. Milojević. Certain Comparative examinations of plane geometries according to Cayley-Klein. *Novi Sad J.Math., Vol. 29, No. 3*, 159–167, 1999.
- [8] V. Niče. Uvod u sintetičku geometriju. *Školska knjiga, Zagreb*, 1965.

ABOUT THE AUTHORS

1. Ana Sliepčević is retired Associate Professor. Her research interest are projective geometry, Euclidean and non-Euclidean planes treated by synthetic methods. She can be reached by e-mail: anas@grad.hr
2. Ivana Božić is Assistant Professor at the Polytechnic of Zagreb, Department for Civil Engineering. She can be reached by e-mail: ivana.bozic@tvz.hr.

ANALYSIS FOR ACCESSIBILITY TO THE CONVENIENT FACILITIES FROM THE RESIDENCES

Ai SAKAKI and Ryotaro KOBAYASHI
Setsunan University, Japan

ABSTRACT: Quality of housing has been improved in recent years, and demands on the living environment have been diversified in Japan. In this study, focusing on the importance of convenience in living environments, accessibility to convenient facilities from the residences was analyzed. This study aimed at clarifying problems with road networks and placement of the facilities, and improving convenience in residential areas. The subject area of this study was Neyagawa City in Osaka Prefecture. The subject transportation of this study was cars and walking, and the subject facilities were 59 stations, 651 supermarkets, and 227 post offices. Accessibility was evaluated for each 100m grid. The map of the 100m grid was overlaid on the subject facilities map and the road networks map that had traffic regulation information. The accessibility of each grid was evaluated by the number of facilities that can be reached within 15 minutes. It was found that there were many grids that had excellent accessibility by car, but had poor accessibility by walking. In other words, it was shown that currently-convenient life with a car was not guaranteed to last, in the event of car unavailability due to some reasons (e.g., aging, injury, or disaster) in the future.

Keywords: Accessibility, Residence, Convenient Facilities, GIS, Road Distance.

1. INTRODUCTION

In Japan, the quality of the living environment has been improved in recent years, and demands for the living environment have been diversified. Convenient living environments have attracted great attention. In this study, focusing on the importance of convenience in living environments, accessibility to convenient facilities from the residences will be analyzed. This study aims at clarifying the problems with the road networks and placements of the facilities.

Due to popularization of private vehicles, such as bicycles or cars, individual activity ranges have been widened in present days. Convenient facilities are accessible even when they are distant. But not everyone is able to drive private vehicles. Some of these vehicles require the driver's physical strength, judgment, and a license. Therefore, the users are limited depending on their physical conditions, their

ages and their skills. Even if one currently lives comfortably with a private vehicle, it is possible that convenient living might be discontinued in the future due to some reason.

In this study, focusing on two different types of transportation, the accessibility differences caused by different means of transportation are examined. The subject means of transportation in this study are "walking" and "car". "Walking" is the most basic means, and "car" is a means that can most widen the individual range of activities.

This study defines "accessibility" as a concept to represent the ease of reaching a goal from a starting point. Although a large number of studies have been conducted on accessibility to convenient facilities from residences, little discussion on the means of transportation has been done. The precedent studies evaluated accessibility using the shortest distance and straight-line distance (e.g., Sekine). However, for bicycle or car transportation, the moving

paths are affected by the road network. When the means of transportation is different, the available moving path is also different. When walking, there are moving paths (e.g., a road passing through a public space, and pedestrian roads) that are not on road maps. For car transportation, movements are sometimes limited to a particular direction under the influence of traffic regulations such as one-way only and closed to traffic. However, since the moving speed of cars is faster than that of walking, more distant facilities can be accessible. Therefore, in this study, transportation is simulated considering the difference of moving paths and moving speed, and we attempt to evaluate the accessibility under more realistic conditions.

2. METHOD

2.1 PROCESS

The main subjects of this study and processes are described below. GIS (SIS Map Modeller V7.0 by Informatix Inc.) is used as a tool to create maps and for analysis. The subject area of this study is Neyagawa City in Osaka Prefecture. The subject means of transportation is “car” and “walking”, and the subject facilities are stations, supermarkets, and post offices.

First, we investigate the differences of the total road length that were caused by the differences in the distance measurement method and the consideration of traffic regulation. Then, we clarify the necessity for evaluating the accessibility by not using the straight distance but the road distance with traffic regulation.

Next, we evaluate placement of the facilities by “area coverage rate” and “population coverage rate”. “Area coverage rate” of the subject facilities is defined as a ratio of the areas of the grid with one or more accessible facilities within acceptable time to the area of the entire city. Furthermore, the “population coverage rate” of the subject facilities is defined as a ratio of the population in the grid with one or more accessible facilities within acceptable time to the population of the entire city.

Finally, we evaluate the accessibility by each 100m grid, and discuss the resulting evaluation of the residential area grid.

The brief analytical process of accessibility evaluation is as follows:

- Step 1: A map with 100m grids covering the entire Neyagawa city is created, and the population information in 2010 is added to each grid square.
- Step 2: A map of the subject facilities in Neyagawa City and in the adjacent cities is created.
- Step 3: A map of the road networks for cars and for walking is created.
- Step 4: The map of 100m grids is overlaid on the map of the subject facilities.
- Step 5: Roads that are accessible within an acceptable time from each facility are extracted. An Accessible flag is put on the 100m grid that crosses the road. Accessibility to every facility is evaluated in each grid square.

Figure1 shows sample of the roads accessible within an acceptable time from a facility and the map of 100m grids.

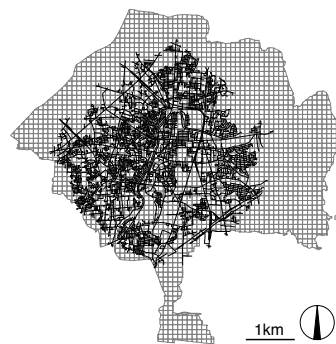


Figure1. Sample of the roads accessible within an acceptable time

2.2 SUBJECT AREA

Figure 2 shows the location of Neyagawa City in Osaka Prefecture that is the subject area in this study.

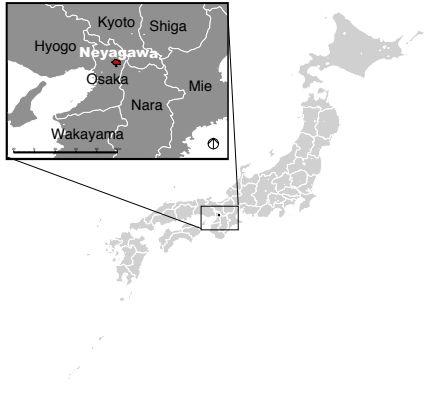


Figure 2 Location of the subject area
- Neyagawa city

Neyagawa city is located in the northeastern part of Osaka Prefecture, 15 km from the center of Osaka city, and 35km from the center of Kyoto city. Therefore, there are many people who commute to Osaka or Kyoto. The population is about 240,000, and the area is about 24.73 km². As for the features of the terrain, Neyagawa city can be divided into two areas. One is the western flat area and the other is the eastern low hills area. The Yodo River flows on the northwest part of the city. There are various residential areas. For example, a densely populated district with old wooden buildings, a detached house district, and a large public rental housing district.

In this study, accessibility is evaluated, taking into account these characteristics of Neyagawa city.

2.3 SUBJECT FACILITIES

Subject facilities that are essential for daily living are chosen, regardless of age, sex, and family structure. "Station" as a public transportation facility to assist broad areas of transportation, such as commuting, "supermarket" is chosen as a comprehensive facility for purchasing daily necessities and "post office" as a facility that provides nationwide services that are required for living, such as postal mail, savings, and insurance. There are 637 facilities (59 stations, 651 supermarkets and 227 post offices) in Neyagawa city and in the adjacent cities. The internet service "i Town Page" is used for collecting the location information of all facilities, and GIS is used to create the map of the subject facilities. Figure 3 shows the map of the subject facilities.

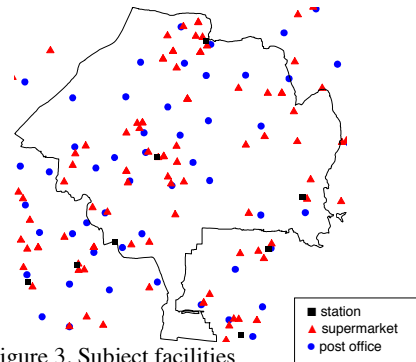


Figure 3. Subject facilities

2.4 TRANSPORTATION CONDITION

The subject means of transportation in this study are "walking" and "car". "Walking" is the most basic means, and "car" is the means that can widen the individual range of activities the most.

Acceptable travel time by each means is discussed. It is desirable that the frequently-used facilities for living are easily accessible. We focus on the acceptable time to access the convenient facilities. Referring to the documentation and previous research on the effec-

tive distance of each facility, the acceptable time is set to be 15 minutes.

Then transportation speed for each means is discussed. In this study, car transportation speed is set to be 20.3km/h, and walking is 4.8km/h, referring to the previous studies and documents. Since cars require time for parking, 7 minutes of parking time was set.

According to the acceptable travel time and transportation speed, the acceptable travel distance (movable distance within the acceptable travel time) for each means is calculated. For car transportation, parking requires seven minutes and traveling takes up the remaining 8 minutes at a speed of 23km/h. Therefore, the acceptable travel distance will be 2700m. For walking, traveling takes 15 minutes at a speed of 4.8km/h, and the acceptable travel distance will be 1200m. Car transportation allows about two times longer travel distance than walking, even taking into account parking time.

2.5 ROAD NETWORK DATA

“Mapple Road Network Data” by Shobunsha Publications Inc. is used as road network data for car transportation, and “Mapple Walking Network Data” by Shobunsha Publications Inc. is used for walking transportation.

Figure 4 shows the road network data. For car transportation, accessibility is evaluated for car-accessible roads taking into consideration three types of traffic regulation information: “no regulation”, “one way”, and “no through road”. “No through road” includes a road that is not allowed to pass through due to the traffic regulation, and that is physically limited due to its narrow width. For walking, accessibility is evaluated for open roads excluding exclusive highways, as well as for roads for pedestrians including pedestrian bridges, underground passages, road passing through a public space, and pedestrian roads.

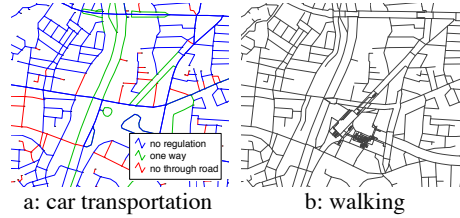


Figure4. Road network data

3. RESULT

3.1 RESULTS RELATED TO COMPARISON OF TOTAL ROAD LENGTH

To confirm the necessity for evaluating the accessibility using the road distance with traffic regulation, we investigated four stations in Neyagawa city. We chose roads within 2700m from each of four stations in the following three methods, and calculated the total road length.

- Method A: To choose the roads within 2700m straight-line distance.
- Method B: To choose the roads within 2700m road distance without traffic regulation.
- Method C: To choose the roads within 2700m road distance with traffic regulation.

The bigger the difference between method A and B is, the higher the necessity of evaluating the accessibility using the road distance. The bigger the difference between method B and C is, the higher the necessity of evaluating the accessibility taking into consideration the traffic regulation.

Table 1 shows the result of the comparison of the total road length in Methods A – C.

Table 1. Comparison of the total road length in Methods A - C

Distance measurement method	Method A	Method B	Method C
	straight-line distance	road distance	
Traffic regulation	-	no traffic regulation	with traffic regulation
Neyagawa city	632.4km (100%)	419.5km (66%)	338.8km (54%)
Korien	541.6km (100%)	391.0km (72%)	335.0km (62%)
Kayashima	691.0km (100%)	476.1km (69%)	335.7km (49%)
Higashi-Neyagawa	455.7km (100%)	292.9km (64%)	225.4km (49%)

When the total road length of method A is set to 100%, that of method B decreased to 64% - 72%, and the one of method C decreased to 49 - 62%. From the above result, it was found that evaluation using straight-line distance was twice the amount compared with the evaluation using road distance and traffic regulation. There are many car-inaccessible narrow roads around Kayashima station. This caused the significant decrease in the total road length when considering traffic regulation.

For comparing the characteristics of the accessible range shape in method B - C, a convex polygon was created from the extracted roads, and the degree of circularity was calculated. The result is shown in Figure 5. Numbers in this figure show the degree of circularity and the area of the convex polygon in Method C.

About Higashi-Neyagawa station, the degree of circularity was smaller than others. This was caused by the small accessible area to the southeast, because there were few roads in the hilly area to the southeast. This caused the significant decrease in the total road length when using road distance. With respect to Neyagawashi station and Korien station, the shape was almost a circle. But this was influenced by the Yodo River and the accessible area to the northwest was smaller.

From the above, when a subject area has many narrow roads, evaluation that takes into consideration traffic regulation is necessary, and when a subject area is influenced by the terrain, evaluation based on road distance is

necessary. For evaluating the accessibility under more realistic conditions, the necessity of evaluating the accessibility based on road distance with traffic regulation is clarified.

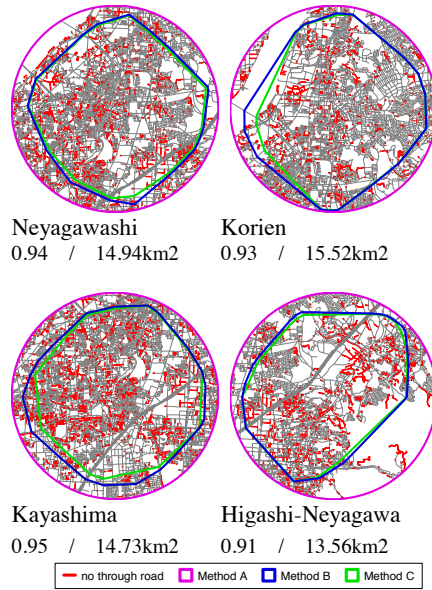


Figure 5. Comparison of the convex polygon in Method B - C

3.2 RESULTS RELATED TO POPULATION DISTRIBUTION

Figure 6 shows a thematic map of population density (persons/km²) in Neyagawa city, 2010. The population concentrates along the Keihan line running north-south, and gradually reduces towards the northwestern part along the Yodo River, and hilly southeastern part of the city. The population density of Neyagawa city is about 10,000 persons/km². Therefore, in relative comparison within Neyagawa city, 100m grids with 10,000 persons/km² or higher value can be said to have a higher population density. Figure 7 shows the areas with a relatively high population density in Neyagawa city. The accessibility of the areas will be mainly discussed in the following evaluation of accessibility.

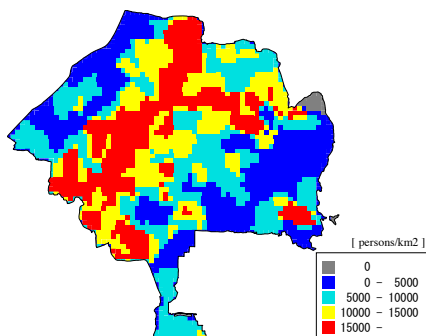


Figure 6. Population of Neyagawa city (2010)

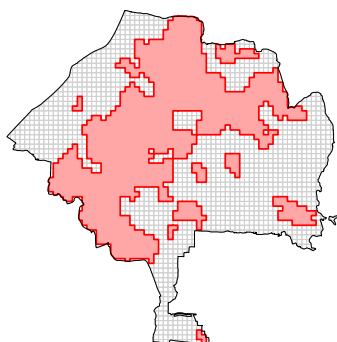


Figure 7. Areas with a relatively high population density in Neyagawa city

3.3 EVALUATION OF STATION ACCESSIBILITY

Figure 8 shows the locations of stations, and figure 9 shows the numbers of accessible stations within 15 minutes. The area coverage rates of the stations are 84% for car transportation, and 43% for walking. The population coverage rates are 92% for car transportation, and 47% for walking.

The coverage rates for walking are lower than 50%, showing the presence of inconvenient areas when assuming walking transportation. Specifically, due to the long distance between the Korien station and the Neyagawashi station, there are inaccessible areas to the stations within 15 minutes, even though they are

located along the Keihan railway. According to some materials, there used to be a station between the Korien station and the Neyagawashi station. The station was closed when a playfields park nearby was closed and the situation remained the same up until today. It is revealed that even though many residents currently live in the area, there are widespread areas without station accessibility within a 15 minute-walk, due to the long distance between the stations, and there is room for discussion.

Furthermore, the northwestern part is a very inconvenient area that has only a few accessible stations even for car transportation. It is partly due to the Yodo River that runs northwest of the area and restricts people's movement to the north and west directions, as well as the long distance from the Keihan line running in a south-east direction.

Figure 10 shows the areas accessible within 15 minutes by car, but inaccessible by walking. About 43% of the total area has no problem for car transportation, but is inconvenient for walking. Many commuters to outside of the city live in Neyagawa city; however, parking lots are not sufficiently facilitated around the stations. Therefore, private car transportation to the stations is unlikely, and it is thought that the importance of the station accessibility from the residential areas is quite high in Neyagawa city. Bus transportation is perhaps one of the public transportation systems that complements the high ratio of the poor station accessibility areas without cars. Evaluation of accessibility to the station without cars after adding buses to a subject facility is a future task.

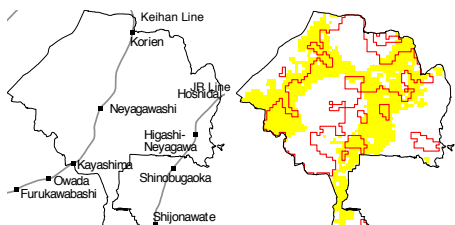


Figure 8.
Locations of stations

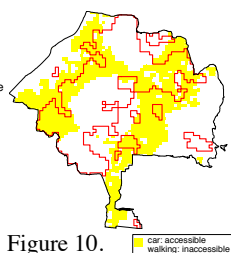
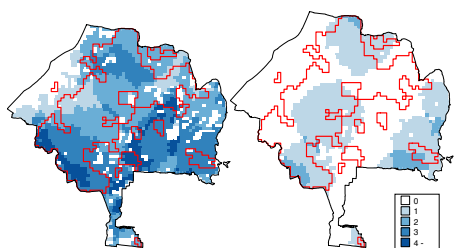


Figure 10.
Areas accessible to stations by car but inaccessible by walking

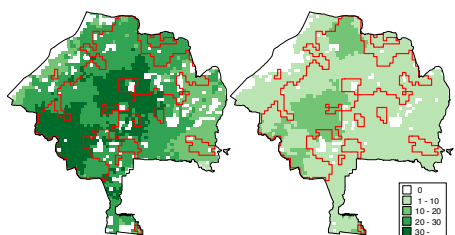


a: car transportation
b: walking
Figure 9. Numbers of accessible stations

3.4 EVALUATION OF SUPERMARKET ACCESSIBILITY

Figure 11 shows the numbers of accessible supermarkets within 15 minutes. The area coverage rates of the supermarkets are 88% for car transportation, and 90% for walking. The population coverage rates are 94% for car transportation, and 92% for walking.

It can be said that most areas in the city have excellent accessibility to the supermarkets, excluding a part of the northwestern area.

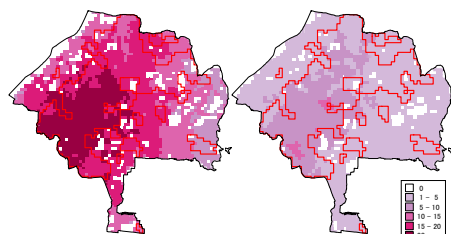


a: car transportation
b: walking
Figure 11. Numbers of accessible supermarkets

3.5 EVALUATION OF POST OFFICE ACCESSIBILITY

Figure 12 shows the numbers of accessible post offices within 15 minutes. The area coverage rates of the post offices are 88% for car transportation, and 92% for walking. The population coverage rates are 94% for car transportation, and 95% for walking.

Among the three types of facilities, accessibility to the post offices is the highest maintained over the entire city regardless of type of transportation. It can be evaluated that the post office accessibility is ensured even in the northwestern part, which has poor evaluations for other facilities. The number of facilities are about one third of the supermarkets; however, the area coverage rate and the population coverage rate are both about the same or higher than those of the supermarkets. This may be partly because the post offices were originally public facilities and their locations were chosen by emphasizing equity. They still have a role to widely support living of the residents, even after they are privatized.



a: car transportation
b: walking
Figure 12. Numbers of accessible post offices

3.6 OVER-ALL EVALUATION OF ACCESSIBILITY TO THE THREE TYPES OF FACILITIES

In the preceding paragraph, areas with accessibility difficulties were discussed for three types of facilities: stations, supermarkets, and post offices. The results are shown in Table 2.

Table 2. Area coverage rates and population coverage rates of three types of facilities

	area coverage rates		population coverage rates	
	car transportation	walking	car transportation	walking
station	84%	43%	92%	47%
supermarket	88%	90%	94%	92%
post office	88%	92%	94%	95%

However, considering convenience, the minimum requirement is an accessibility within 15 minutes to at least one of each type of three subject convenient facilities. It can be said that as the number of accessible facilities rises, convenience improves.

The over-all accessibility of each grid is evaluated based on the following conditions:

Condition 1: whether it is accessible to at least one of each type of three subject convenient facilities

Condition 2: whether the total number of accessible facilities is relatively high in Neyagawa city

For the condition 2, the numbers of accessible facilities for all grids in Neyagawa city are first determined. Compared with the median value, the number of the accessible facilities of each grid is evaluated.

According to conditions 1 and 2, the grids are comprehensively evaluated in four phases shown in table 3.

Table 3. Over-all evaluation criteria

Condition 1 \ Condition 2		Accessible to at least one of each type of three subject convenient facilities	
		Accessible	Inaccessible
Total number of accessible facilities is greater or fewer than the median	Greater	Good	Somewhat poor
	Fewer	Somewhat good	Poor

Figure 13 shows the result of the four-phase evaluation. By comparing the figures, there are many grids that are highly evaluated for cars, but poorly evaluated for walking. In other words, it is shown that currently-convenient life with a car is not guaranteed to last, in the event of car unavailability due to some reasons (e.g., aging, injury, or disaster) in the future.

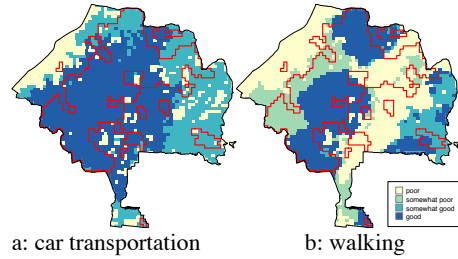


Figure 13. Evaluation map of accessibility to all facilities

4. CONCLUSIONS

In this study, accessibility to convenient facilities (stations, supermarkets, and post offices) was evaluated taking into consideration traffic regulation information and means of transportation in Neyagawa city. The areas with accessibility problems to convenient facilities were clarified for the entire city.

It was found that evaluation using straight-line distance was twice the amount compared with the evaluation using road distance and traffic regulation. When a subject area has many narrow roads, evaluation that takes into consideration traffic regulation is necessary, and when a subject area is influenced by the terrain, evaluation based on road distance is necessary. For evaluating the accessibility under more realistic conditions, the necessity of evaluating the accessibility based on road distance with traffic regulation is clarified.

For all three types of facilities, there are inaccessible areas within 15 minutes. However, the population coverage rate of post offices and that of supermarkets exceed 90% for both cars

and walking, and their accessibility can be highly rated. Poor walking accessibility to stations, whose number is smaller, was widely found. Evaluation of bus usage to possibly solve such inconveniences is a future task.

The northwestern part of the city is located on the left bank of the Yodo River. In order to cross the river, a huge detour to a bridge is necessary. For this reason, while movements in the center of the city are relatively flexible in 360 degrees, movements in the northwestern area towards the Yodo River direction are restricted. Therefore, regardless of the types of the facilities and means of transportation, areas with poor evaluations were widely found.

The subject facilities of this study included only three types of convenient facilities. But there also are other facilities that are necessary for a comfortable living environment, such as hospitals, parks, elementary schools, etc. Furthermore, this study only examined car transportation and walking as main transportation means; however, there are more transportation means to be studied including buses and bicycles. Addition of more subject facilities and transportation means, and conducting a more detailed evaluation in the future, may help improvement of the living environment.

REFERENCES

- [1] H. Kohsaka, T. Sekine. Social and Economic Spatial Analysis Using GIS, Kokon Shoin Inc, 2005
- [2] NTT DIRECTORY SERVICES CO. “iTownpage”, <http://itp.ne.jp/?rf=1> , [Accessed July 2014]
- [3] T. Sekine. Geographical Studies of Quality of Life and Living Environment: A Review and Perspective. Annual report of economic geographers, 39(3), 221-238, 1993.

ABOUT THE AUTHORS

1. Ai SAKAKI, Dr. Eng., is an assistant professor of Department of Living and Environmental Design, Faculty of Science and Engineering, Setsunan University. Her research interests are Graphic Science and Geographic Information System. She can be reached by e-mail: sakaki@led.setsunan.ac.jp.
2. Ryotaro Koba is a graduate of Department of Living and Environmental Design, Faculty of Science and Engineering, Setsunan University. His research interests are Graphic Science and Geographic Information System.

ANALYSIS ON SEQUENCE OF ARCHITECTURAL SPACE BY USING VR WALK-THROUGH SYSTEM

Kensuke YASUFUKU
Osaka University, Japan

ABSTRACT: This study evaluates architectural space in terms of a sequence. In this paper, we propose a computational analysis method of a visual sequence along a walking path by using a VR walk-through system. A display image of our system is shown from a first-person perspective. We consider a perspective to be the visual field of the users. To convert a perspective into the visual field, we take advantage of depth buffer data, which can specify how far a pixel of an object is on the perspective. A virtual reality display, which employs a wide range view angle and a head-tracking system, also enables us to compare the relation between a sequence and head motion. To analyze a sequence of visual fields, we introduce evaluation indicators based on the isovist theory. As a result, the changes of the evaluation indicators and the ratio of those indicators evaluate the characteristics of sequences quantitatively.

Keywords: Architectural space analysis, Virtual reality, Walk-through, Isovist.

1. INTRODUCTION

Architectural space is experienced by human body walking through it. The continuous changes of the visual fields along a walking path are called sequence, but conventional architectural drawings cannot represent such sequence. So far, the descriptions of sequence, such as Notation [1], Motation [2], have been proposed. Those are diagrammatic representations of movements through time and space like traditional musical scores. However, to imagine the sequence from a description, a particular skill is required. We have investigated a new representation method of a visual sequence on a walk-through system [3]. We also proposed a computational analysis method of a visual field based on the isovist theory [4].

To provide more immersive experience on a walk-through system, virtual reality (VR) displays are effective. With such displays, it becomes possible to simulate visual sequence of architectural space based on human perception. In recent years, low-cost but immersive head mounted displays (HMDs) have become availa-

ble because small high-definition displays are becoming common and software technologies in computer graphics are able to correct a distorted image in real time. In this paper, we propose an analysis method of a visual sequence of architectural space along a walking path by using a VR walk-through system with a HMD.

2. METHOD

2.1 VR walk-through system

A VR walk-through system displays a walk-through animation image in a virtual 3D building on a VR display and enables users to look around interactively. The programming language for our developed VR walk-through system is C++. The 3D graphics library is OpenGL.

2.2 VR display

VR displays provide users with a high level of immersion by covering their entire visual field. A CAVE system [5] is a typical VR display device consisting of a cube-shaped room in which the walls are rear-projection screens. A HMD is

also a virtual reality display device that has a small display optic in the front of each eye. Though we have developed the walk-through system on a CAVE system, the issues that researchers are faced with are size, portability and cost [6]. In recent years, low-cost but immersive HMDs have become available because small high-definition displays are becoming common and software technologies in computer graphics are able to correct a distorted image in real time. In this study, we use an Oculus Rift™, which is a HMD developed for immersive video game play [7].

The specifications of the Oculus Rift are as follows. Stereoscopic video images display a 7 inch LCD panel. The resolution is 1280 x 800, which is split between both eyes, yielding 640 x 800 per eye. Dual optical zoom lenses positioned over the eyes provide a 90 degree horizontal and a 110 degree diagonal stereoscopic 3D perspective. The Oculus Rift hardware includes a gyroscope, accelerometer, and magnetometer. These sensors determine the orientation of the user's head in the real world, and synchronize the user's virtual perspective image over all 360 degrees.

2.3 Stereoscopic image creation

The Oculus Rift is designed for video games with a first person view. To integrate the Oculus Rift with original applications, the Oculus SDK (software development kit) is available. The Oculus SDK includes the specification for rendering stereoscopic images and APIs for fetching the sensor data. To create stereoscopic images on the Oculus Rift, a 3D scene needs to be rendered in the split screen, where each half of the screen is used for each eye. Unlike stereo TVs, the Oculus Rift does not require off-axis or asymmetric projection. Instead, the projection axes are parallel to each other, as shown on the right-hand side of Figure 1. To integrate the Oculus Rift with our walk-through system, we added a stereo rendering process based on the Oculus Rift specification.

To create stereoscopic images, our VR

walk-through system renders a 3D building twice. A point of view is set at the position of the right and left eyes in virtual space. The left eye image is rendered on the left half of the screen and the right eye image is rendered on the right hand of the screen. The display of the Oculus Rift is 93.6 mm high and 149.7 mm wide. The distance between the eyes and the display is 41 mm. The distance between the pupils is assumed to be 64 mm. The gap of 5.44 mm between the center of the picture plane and the center of vision is also to be considered to calculate the perspective projection matrix. The positional relation between the eye position and the picture represents an approximately 84.8 degrees horizontal and an approximately 97.6 degrees vertical perspective. In practice the dual optical zoom lens broadens the horizontal perspective to 90 degrees.

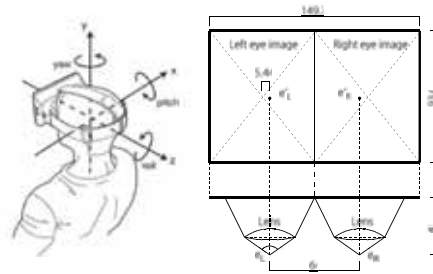


Figure 1: Perspective projection on the Oculus Rift.

2.4 Distortion correction

The optical zoom lens on the Oculus Rift magnifies the image to provide an increased view angle and to focus on the screen, but this comes at the expense of creating a pincushion distortion of the image. This radially symmetric distortion can be corrected in the software by adding a barrel distortion that cancels it out. This distortion correction model is expressed in terms of the distance r from the center of vision. It is helpful to regard a point in the image in polar coordinates (r, θ) for a point in the image. The distortion correction model per-

forms the transformation in which the scaling function is given based on the parameter of the Oculus SDK.

For such image processing, the VR walk-through system renders a scene with the off-screen texture having the same dimensions as the screen. Then, the off-screen texture is applied to the post-processing filter of the barrel distortion by using the correction expression. Finally, the off-screen texture is applied to a half-screen quad for an eye and displayed on the screen. However, the barrel distortion makes the image smaller. For this reason, the system renders an image larger than the size of the screen. Current GPU technology enables this graphics processing in real time by using a fragment shader technique of OpenGL. Figure 2 shows the barrel distortion image on our VR walk-through system to cancel a pincushion distortion.

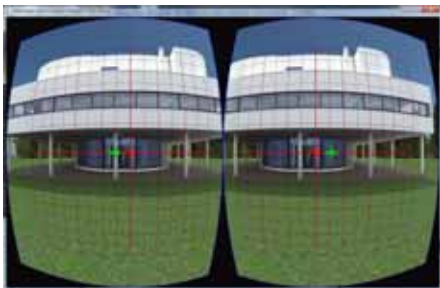


Figure 2: Barrel distortion image to cancel pincushion distortion.

2.5 Head tracking

The Oculus Rift hardware includes a gyroscope, accelerometer, and magnetometer. We combine the information from these sensors through the Oculus Rift SDK to determine the orientation of the user's head in the real world, and to synchronize the user's virtual perspective in real time. The Oculus Rift orientation is a rotation angle in the right-hand coordinate system. The rotation angle is maintained as a unit quaternion, but can also be reported in yaw-pitch-roll form. Our VR walk-through system adds the

yaw-pitch-roll to the camera rotation angle. In a general walk-through system, a user moves forward, backward, left, and right by using a keyboard, and looks left, right, up, and down by using a mouse. By using the head rotation data, a user can look around instead of using a mouse.

2.6 Reconstruction of 3D visual field

In any discussion of vision it is necessary to distinguish between the retinal image and what man perceives. Gibson has technically labeled the former the visual field and the latter the visual world [8]. In this paper, we focus on architectural space based on visual world constructed by using a sequence of visual field.

A walk-through image of our system is shown from the first-person perspective. We consider the perspective to be the visual field of the users. To convert the perspective into a visual field, we take advantage of depth buffer data, which can specify how far a pixel of an object is on the perspective. To analyze a visual field, we quantify the geometric characteristics based on isovist theory. Our system especially focuses on the 3D visual field within the view angle, whereas the conventional isovist is the set of all points visible in 2D space from any given vantage point. Consequently, the visible area in our system is different from an isovist because our system considers the view direction and the view angle in 3D space. Therefore, our system has the advantage of simulating a more real visual experience.

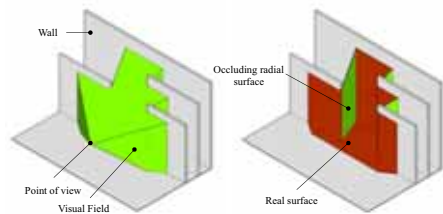


Figure 3: 3D visual field.

Figure 1 shows the definition of terms relating to a 3D visual field. A visual field is cut according to the shape of the environment and the point of view. The boundary of the visual field is decomposed into three parts: real surfaces, occluding radial surfaces, and region-boundary surfaces. Real surfaces, such as walls, floors, ceilings, etc., are visible. Occluding radial surfaces represent the depth to which environmental surfaces partially cover each other, as seen from the point of view.

2.7 Evaluation indicators of 3D visual field

As evaluation indicators of a 3D visual field, our system calculates the visual volume, the real-surface perimeter and the occlusivity. The view volume V is that of the visual field. The real-surface perimeter P of the visual field is the sum of the real surfaces. The occlusivity Q of the visual field is the sum of the occluding radial surfaces. In addition, our system separately calculates the indicators of each element, such as walls, floors, ceilings, as opposed to a conventional method dealing with each element as the same environmental surfaces. The type of a surface can be arbitrarily set by assigning a material in the 3d modeling phase. Specifically, V_x represents the volume of the visual field cut by element x . P_x represents the real-surface perimeter of element x . Q_x represents the occlusivity covered by element x .

In isovist theory, isovist fields are also defined by a scalar field at each point. Our system cannot calculate these indices at each point because the indices depend on not only the view point but also the view direction. Instead of displaying isovist fields, our system displays monitoring graphs of these indices.

2.8 Case study

We apply our VR walk-through system to a real building to evaluate the indicators for analyzing the changes of the visual field along a walking path. The indicators are the visual volume, the real-surface perimeter and the occlusivity of the visual field. The subject building is the Villa

Savoye designed by Le Corbusier. The Villa Savoye represents not only a demonstration of Le Corbusier's five points of a new architecture but also an architectural promenade that embodies the concept that an architectural experience is the sequence of movement. Therefore, the subject building is appropriate to evaluate architectural sequence quantitatively.

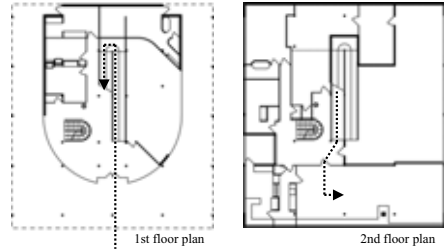


Figure 4: Case study plan.

We created the 3D model and set a walking path on which a user enters the building and walks on a ramp up to the salon, as shown in Figure 4. The subject is one user this time because the primary aim of this case study is to verify performance characteristics of our VR walk-through system. The user has already learned how to operate the system in advance. When the user wears the headset and watches the walk-through animation along the walking path, he can look around freely. To analyse the visual field, we use a single perspective image, while the user looks at the stereoscopic images. For system configurations, the horizontal view angle is 90 degrees per eye, the maximum visual distance is 50 m and the resolution of the display is 1280 x 800.

3. RESULTS AND DISCUSSIONS

3.1 Head motion along the walking path

Though the walking path of the case study is set in advance, a user wearing the headset can look at 360 degree space in accordance with his head motion. Figure 5 shows the head motion data of the user and the walk-through images.

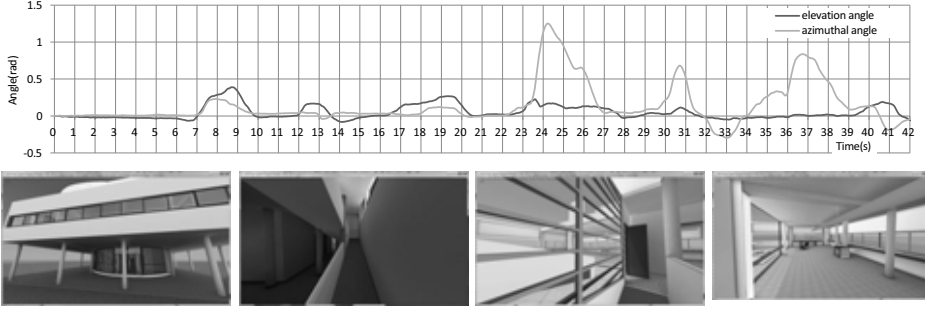


Figure 5: Head motion data and walk-through images.

As shown in the figure, when focusing on the elevation angle, the user looked up from 7 seconds to 9 seconds after starting on the walking path. Before the user entered the pilotis space, he looked up at the second-story walls from the exterior space. From 12 seconds to 14 seconds, when entering the building, the user looked up. It is assumed that the user checked the doorway height and confirmed it was high enough to pass through without hitting his head. Then, the user looked up from 16.5 seconds to 19.5 seconds and from 23 seconds to 27 seconds as he walked up the ramp. When the second story came into the sight, the elevation angle was becoming horizontal.

In terms of the azimuthal angle, from 24 seconds to 26 seconds after the start, the user turned his head before changing the direction along the walking path at the landing area of the ramp. While the user walked up the ramp, the roof garden became visible through the left side of a window and the user turned his head to that direction. After 35 seconds, the data shows that the user entered the salon and looked around.

From the above, we confirmed the head motion characteristics, in which the user's head turned to the direction where he wanted to go and look. To analyze the visual field similar to human perception, it is important to take these head motions into consideration.

3.2 Analysis of sequence in the visual field

Figure 6 shows the changes of evaluation indi-

cators along the walking path by considering the head motion in the case study. The evaluation indicators were calculated for each building element: walls (V_1 , P_1 , Q_1), floors (V_2 , P_2 , Q_2), ceilings (V_3 , P_3 , Q_3), columns (V_4 , P_4 , Q_4), and window frames (V_5 , P_5 , Q_5).

In terms of the visual field shielded by the walls, the trends of changes of the view volume (V_i) and the real-surface perimeter (P_i) are almost the same. This can be explained as follows. Both V_i and P_i increase if the area of a wall in a visual field becomes large. However, the ratio of V_i to P_i is different according to the scene. The ratio depends on the positional relationship of the view point and walls. Specifically, if the distance to a wall is long, the view volume is increased. In addition, the view volume of a front wall is bigger than that of a side wall in the depth direction, even though the areas of those walls are the same. It is likely that the proportion or direction of walls in a visual field can be analyzed quantitatively by using the ratio of V_i and P_i .

For example, when the user was outside and looked at the second-story wall and the rooftop wall from the front, the ratio was more than 2.0. On the other hand, when the user entered the building and walked up the ramp, the ratio was less than 1.0 because of the surrounding side walls in the depth direction. After the user entered the salon, the ratio was more than 2.0 again because of the open space. Therefore, the ratio of P_i and V_i probably evaluates the open-

ness and direction of a visual sequence.

Next, we focus on the visual field shielded by the floors. When the user was outside at the start of the walking path, the user saw part of the floor through the front door and the windows. As the user approached the building, the real-surface perimeter P_2 increased. On the other hand, as the distance between the view position and the floor became shorter, the view volume V_2 was unchanged. Inside the building, the change trends and the ratio between V_2 and P_2 are almost the same because the floors are horizontal surfaces.

Next, we focus on the visual field shielded by the ceilings. When the user was outside at the start of the walking path, the user looked around the pilotis ceiling into the distance. The

real-surface perimeter P_3 is very large and larger than the view volume V_3 because the ceiling is horizontal. Inside the building, the change trends and the ratio between V_3 and P_3 are almost the same because the ceilings are horizontal surfaces.

In terms of the visual field shielded by the columns, the occlusivity Q_4 increased when the user saw the pilotis columns. If there is no wall between columns, Q_4 is relatively large comparing V_4 and P_4 .

Finally, we describe the visual field shielded by the window frames. When the user was outside, the occlusivity Q_5 was not so high. On the other hand, when the user saw the external landscape through the window frames of the building when he was on the ramp and in the

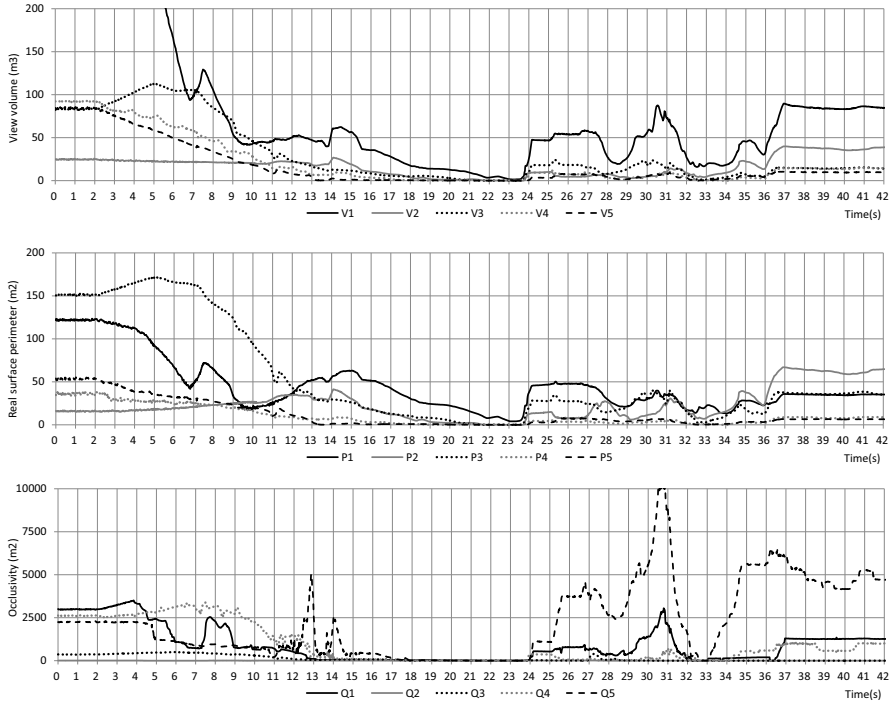


Figure 6: Changes of the evaluation indicators.

salon, Q_5 was a high value. In the salon especially, no big changes occurred in the Q_5 value, even if the user looked around. It is assumed that the horizontal windows, which are one of the five points of the new architecture, maintain the occlusivity.

4. CONCLUSIONS

We proposed a computational analysis method of a visual sequence of architectural space along a walking path by using a VR walk-through system. The virtual reality display, which has a wide range view angle and a head-tracking system, enabled us to compare the relation between visual field and head motion. To analyze a sequence of architectural space similar to human perception, it was important to take head motions into consideration. We also quantify the geometric characteristics of visual field based on isovist theory. As evaluation indicators, our system calculates the visual volume, the real-surface perimeter and the occlusivity of visual field. As a result, the changes of evaluation indicators and the ratio of those indicators evaluated the characteristics of sequence along a walking path quantitatively.

Further research is needed to provide a method for clarification of characteristics of sequence based on the proposed indicators. Thereby it is possible to analyze and evaluate architectural space considering human perception.

ACKNOWLEDGMENTS

This work was supported by JSPS KAKENHI Grant Number 24760490.

REFERENCES

- [1] P. Thiel. *People, Paths and Purposes: Notations for a Participatory Envirotecture*. Univ. of Washington Press, 1997
- [2] L. Halprin. *The RSVP Cycles: Creative Processes in the Human Environment*. George Braziller, 1970
- [3] K. Yasufuku, Y. Deki, H. Abe. Development of a Tool for Analyzing Architectural Space Based on Ambulatory Vision. *Proceedings of the 15th International Conference on Geometry and Graphics*, 2012
- [4] Benedikt M. L. To take hold of space: isovist and isovist fields. *Environment and Planning B*, Vol.6, pages 47-65, 1979
- [5] C. Cruz-Neira, D. Sandin, and T. DeFanti. Surround-Screen Projection-based Virtual Reality: The Design and Implementation of the CAVE. *SIGGRAPH'93: Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques*, pages 135-142, 1993
- [6] K. Yasufuku. Visualization of Evacuation Simulator on a High-definition Immersive Projection Display. *Proceedings of the 13th International Conference on Geometry and Graphics*, 2008
- [7] Oculus VR: 2013. Oculus Rift web site. Available from: Open Source Repository <<http://www.oculusvr.com/>> (accessed 3 June 2014).
- [8] J. J. Gibson. *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin., 1980

ABOUT THE AUTHOR

1. Kensuke YASUFUKU, Dr. Eng., is an assistant professor at Osaka University. His e-mail and postal address are as follows: yasufuku@cmc.osaka-u.ac.jp, Cybermedia Center, Osaka University, 5-1 Mihogaoka, Ibaraki-shi, Osaka 567-0047, Japan

APPLICATION OF CURVES AND SURFACES OF HIGHER ORDERS OBTAINED BY INVERSION IN THE PRACTICE OF ARCHITECTURE

Gordana ĐUKANOVIĆ, Đorđe ĐORĐEVIĆ, Marija Đ. OBRADOVIĆ and Slobodan MIŠIĆ

The University of Belgrade, Serbia

ABSTRACT: Throughout the world, teams of architects are constantly competing to find the most attractive form of future buildings that apart from being functional must meet the criterion of aesthetic beauty. There are numerous architectural objects that are thought to be works of art and their beauty and shape are not only the pride of their authors, they are symbols of the cities and states in which they were built. In order to make the realization of the designed objects possible, geometric shapes (curves and surfaces) must have clearly defined geometry and they must satisfy a number of structural requirements. The principles of their design and construction define the purpose to which they will be used and determine their usability. The major problem that designers face is how to set the mechanism for the design of surfaces and their contours which are curves of different orders

The advance of digital technology in the last twenty years has led to formal freedom in the design of buildings and in their representation in the virtual space. The combination of non-standard geometry and CAD tools has enabled us to express and realize architectural structures in new and modern ways. Transformation of doubly curved surfaces into structural and physical objects always entails huge costs and serious problems both of geometric or static nature.

The paper explains the ways surfaces of higher orders are constructed and used in the practice of architecture. The basic transformation is inversion which is interpreted in two ways: as quadratic transformation in the classical projective geometry and as pure symmetry in the relativistic geometry, with constant comparison of the two geometric systems and their opportunities for explanations and generalizations.

NURBS curves are used in the computer graphics to represent various forms needed in engineering, design and animation. NURBS or Non - Uniform Rational B -Splines are groups of curves which are due to its flexibility and accuracy used in all processes of modeling, from artwork to production. NURBS curves have precise and well-known definitions. They are studied in mathematics and there is even a whole new field called NURBS geometry, which is studied by mathematicians and programmers. The main advantage of NURBS lies in their ability to provide an accurate presentation of the standard geometric shapes such as circle, ellipse, spherical surface, or torus, as well as representations of free-forms such as forms of industrial products or a human figure. The surfaces obtained by inversion consist of (one, two, or three) circular cross-sections and therefore they have certain advantages both in the design and in the manufacturing phase compared to the surfaces designed using NURBS geometry.

We cannot talk about the application of various surfaces in architecture without mentioning diagrid. Diagrid is a structural system whose diagonal members are connected in such a way that they act like a weaving or knitting network which can 'wrap' around various geometric forms whether they are obtained by the inversion, or in some other way. Diagrid a structural system that has been widely used for newly constructed high-rise buildings made of steel and it consists of triangular structures

with diagonal support beams.

Keywords: Inversion, Axial symmetry, Curves of the 3rd and 4th order, Surfaces of the 3rd and 4th order

1. INTRODUCTION

Construction is the language of architecture, or as Auguste Perret said "Construction is the mother tongue of the architect".



Figure 1: Phoenix International Media Center

Following previous scientific papers (M.Obradovic) [17] the geometry of high-order curves and surfaces obtained by inversion has been compared with the structure of constructed buildings and the possibility of applying these forms in the practice of architecture was analyzed.

One of the best illustrations of the use of high-order surfaces in architecture would be the Phoenix International Media Center (Beijing), which is a surface of the 4th order (Fig.1). This object leaves its mark on the urban landscape of Beijing. Photos were taken from: <http://www.-coolhunting.com/design/phoenix-international-media-center.php>. The building's sculptural shape originates from the "Möbius Strip" (August Ferdinand Möbius, 1790-1868). This surface is of the 3rd order, but the architects have kept only the outline of the strip, because an architectural object cannot be built as a strip. The building was constructed in the diagrid structural system. This system will be described later in the paper. The main directions of the construction elements are calculated as isolines with the degree of leaning that is the most suitable for a diagrid construction. Parametric engineering has become the standard of building design nowadays. Different geometrical forms, such as

surfaces designed using NURBS geometry are feasible. Diagrid, as a new system, seems to be the most suitable for the construction of such buildings.

2. CONSTRUCTION SYSTEMS USED IN THE PRACTICE OF ARCHITECTURE FOR THE SURFACES OBTAINED BY HARMONIC SYMMETRY

The advance of science and technology has permitted the use of complex geometric forms in the practice of civil engineering and architecture. This technological leap has been achieved primarily thanks to the invention of flexible building materials, such as sheet steel used for wall screens and roof cladding on the buildings that are cylindroids in shape. The increasing use of reinforced concrete with pre-stressing presented the major milestone. It allowed the construction of doubly curved surfaces of large spans. Besides these materials, laminated timber is also bendable and as such it can be used for the construction of such surfaces. Plastics and reinforced tarpaulin are among other materials that are used for tents and pneumatic structures. These materials can also be used to design both simple and complex geometric surfaces [18].

Linear construction elements of future structural systems used in the practice of architecture and construction are shaped by the geometry of curves of the second, third, fourth or higher orders [9].

A building can be designed by using parts of surfaces which are then fitted into a whole. They can make complex structural assemblages. In this way, we can make free forms, which are, in practice, still limited by technology and building materials. However, these limitations have become less strict in the recent years with the advent of new materials and new structural systems that are used in ar-

chitecture and construction. An increasing number of buildings are designed in freeform, and many of them (both the finished ones and the ones that are still in projects) win us with their unusual shape and beauty.

Surfaces obtained by harmonic symmetry consist of systems of circular cross-sections that define the surface (one, two, or three). Circular cross-sections are easy to use (Fig. 2 and Fig. 3) [5].

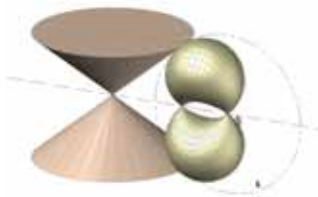


Figure 2: Model of the cone and its harmonic equivalent – spindle cyclide

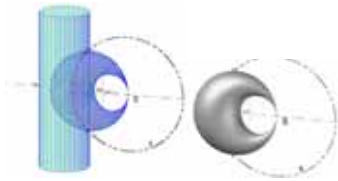


Figure 3: Model of the cylinder and its inverted Dupin Cyclide

On the basis of extensive literature on structural systems the following text provides an overview of structural systems that can use harmonically generated curves and surfaces.

The shell structure has a three-dimensional (stiff) support structure, which carries load primarily with the forces in the plane (uniformly distributed across the thickness of the shell), and sometimes with bending but only in the zone of leaning on or joining the other elements. The shell thickness ranges from less than 1/50 to 1/100 of the whole structure and there is only the direct membrane stress (tension and pressure), without bending moments (or they can be ignored), but only if the shell is uniformly loaded and supported. To achieve a

state that resembles the state of a membrane stress in the shell, the requirements of the membrane theory must be taken into account when selecting the surface or its parts.

With the proper selection of geometry, shells with small thickness can be extremely efficient elements concerning the cost of materials because thanks to their form and reinforcement they have a large carrying capacity despite their thinness. In general, shells can be formed from a variety of surfaces that are characterized by Gaussian curve measure, the product of the curves of the main directions K .

When it comes to the shell structural system, the use of pre-stressed reinforced concrete has produced structures with a longer span, which could not have been even imagined with the use of ordinary concrete.

This paper studies several examples of modern architecture and points to the possibility of their further application. Special attention is focused on doubly curved surfaces obtained by harmonic symmetry - with the positive Gaussian curvature from the harmonic group of cylinders and with the negative Gaussian curvature from the harmonic group of paraboloids. Doubly curved surfaces with the positive Gaussian curvature resulting from the harmonic inversion of the revolving cone - spindle cyclide (Fig. 2) and the revolving cylinder - Dupin cyclide (Fig. 3) are suitable for the shell structural system.

Hanging structures are hung on solid supports installed at great heights, where the main supporting elements are stretched. Supporting elements are small dimensioned, usually steel cables with high tensile strength. This structural system is most commonly used in roofing constructions. The idea of reverse dome has existed since the time of the Incas and the era of ancient China. The surface shown in Fig. 4, which is a harmonic equivalent to the parabolic cylinder, can be used in hanging structures with two systems of pre-stressed high tensile cables[8].

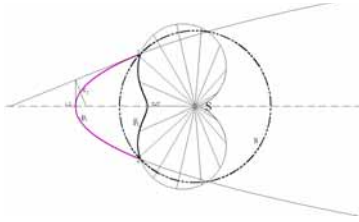


Figure 4: Harmonic equivalent of the parabolic cylinder – the surface of the 4th order with a spike

Harmonic symmetry of the revolving cylinder with the center of inversion S on the contour generating line of the cylinder and the radius of the sphere of inversion - s being the same as the diameter of the circular section of the cylinder produces the 3rd order surface which is suitable for hanging structures (Fig. 5) [13].

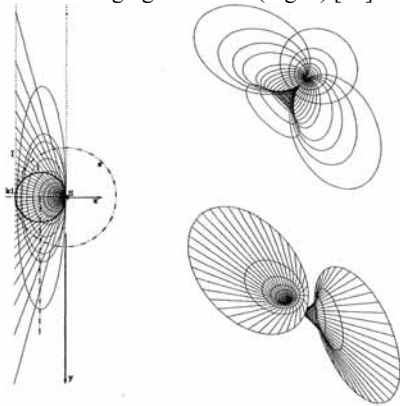


Figure 5: Harmonic equivalent of the revolving cylinder – the surface of the 3rd order

Pneumatic structures – The operation of pneumatic halls and facilities, in general, is based on the vault shaped membrane that is supported by internal air pressure. Pneumatic structures are often said to be airborne. These structures belong to the group of natural structural systems and they are commonly found in zoology and botany. Human muscle tissue and skin are, for instance, stretched under pressure and supported by the solid skeleton.

Pneumatic structures make a special field of textile architecture. Being specific and ingenious in many ways, these structures are in some of their applications simply irreplaceable. They were first used in aeronautics, for the construction of dirigibles and balloons. Although these structures have never come into wide use, they were revived in the form of buildings, during the 80's of the 20th century. Although some of their properties have often been disputed, pneumatic structures are still famed for their quick installation and deinstallation and their low cost.

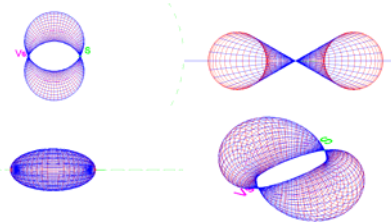


Figure 6: Surface of the 4th order – harmonic equivalent of the revolving cone

Pneumatic structures can use surfaces with two systems of circular cross-sections, which are marked red and blue (Fig. 6). There are three surface projections and one axonometric projection. The main structural feature of this surface is a double curvature with the positive Gaussian curvature (the radii of the double curved surface is on the same side of the tangent plane). A series of spheres can be inscribed into the surface, so that the surface touches the surface of the sphere. This harmonically generated surface satisfies both conditions - circular systems are meridians of individual inscribed spheres [13].

3. APPLICATION OF HIGHER ORDER SURFACES OBTAINED BY HARMONIC SYMMETRY IN LANDSCAPE ARCHITECTURE AND DESIGN OF AMBIENT UNITS

“Cloud Gate” is a public sculpture made by Indian-born British artist Anish Kapoor. The sculpture is located in the Millennium Park in

Chicago in the state of Illinois, USA. It was built between 2004 and 2006 and named “The Bean”, because of its legume-like shape. It is composed of 168 welded plates of stainless steel, and there are no visible seams on its highly polished exterior. The sculpture is popular with tourists because of its unique reflective properties.

This paper analyzes the bean curve shown in Figure 7, obtained by harmonic symmetry of an ellipse with the corresponding position of the center and the circle of inversion.

When the center of inversion S is set beyond the ellipse, its inversion produces a curve of the 4th order – the bean curve[6]. Since the curve has an irregular shape we might think it is asymmetrical. However, it is harmonically symmetrical with the given ellipse and with all the curves that are harmonically symmetrical with the ellipse. This bean curve is centrally symmetrical with respect to the points \bar{C}_1 and $\bar{C}_2 = S$, and it has two circular axis of autosymmetry (\bar{a}_E and \bar{b}_E). The point \bar{S} is an isolated antipodal point of the ellipse, and the point S is an isolated point of the curve of the 4th order[16]. The osculatory circles O_1, O_2, O_3 and O_4 are inscribed into the ellipse apices 1, 2, 3, and 4. These circles are inverted into the osculatory circles of the 4th order curve in the points $\bar{1}, \bar{2}, \bar{3}$ and $\bar{4}$ [7].

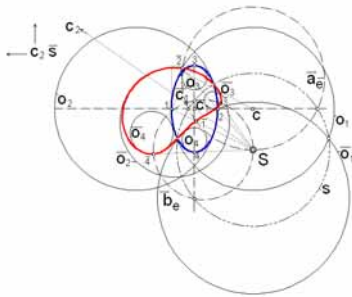


Figure 7: Ellipse and the 4th order curve with two circular axes of harmonic symmetry - bean curve

Photos in Figure 8 and Figure 9 were down-

loaded from the web site: <http://www.anishka-poor.com/484/Studio.html> and they are examples of surfaces that can be obtained by harmonic symmetry. In the photo of the sculpture “Cloud Gate” the surface contour is marked red – it is the bean curve of the 4th order which can be obtained as the harmonic equivalent of the ellipse (Fig.9).



Figure 8: Surface of the 4th order ‘Cloud Gate’ designed in Anish Kapoor’s project and built in Chicago in the U.S.



Figure 9: The bean curve of the 4th order

4. CONSTRUCTION OF BUILDINGS MODELED BY THE GEOMETRY OF THE CURVES AND SURFACES OF THE 3RD AND 4TH ORDER

There is a link between architecture, nature, and science that can’t be broken. The theory of analytical geometry and its practical application on the natural forms in architecture allow us to construct buildings that are not only visually appealing but also admired for their stiffness, hardness and stability. This is best illustrated by objects such as Crescent Moon Tower, designed to be built in Dubai. Some people call it “The Horn”. It is shown in Figure 10 [19]. Figure 11 presents the modeled surface obtained

by harmonic symmetry of the revolving paraboloid, which is, like the building itself, crescent-shaped and has two systems of circular sections.



Figure 10: Moon Tower designed to be built in Dubai



Figure 11: Surface of the 4th order obtained by inversion of the revolving paraboloid

The revolving paraboloid has the z axis. Harmonic symmetry of the paraboloid is done using the sphere s and the center of harmonic symmetry S [14]. The paraboloid is inverted into the surface of the 4th order with a circular axis and circular generating lines, which is shown in Fig.12. The circles which are in the horizontal planes, perpendicular to the circular axis of the rotating quadric (marked as blue rays (for instance - β)), are mapped into circles ($\bar{\beta}$) which are in the planes perpendicular to the circular axis of the equivalent quadric. In the frontal plane of projection, they are also in the ray position and concurrent in the point K . Inversion mapping uses the contour of the paraboloid in the frontal plane of projection, circles

in H-plane (β) and axis a which is in F-plane. The axis of the cone a (z axis) is mapped into the circular axis that passes the point S . [15]

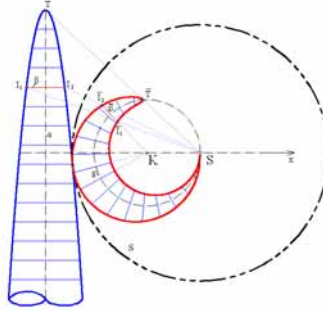


Figure 12: Mapping the revolving paraboloid into the crescent-shaped surface of the 4th order

By selecting different spheres- s and centers of harmonic symmetry S , we can produce surfaces of attractive shapes that are suitable for constructions because of the system of circular cross-sections that are obtained by the above described mapping. To construct buildings modeled in this way, in a system such as Diagrid, it is necessary to find optimal directions of isolines, which would represent the directions of the supporting beams of the diagrid structural system and optimize its capacity.

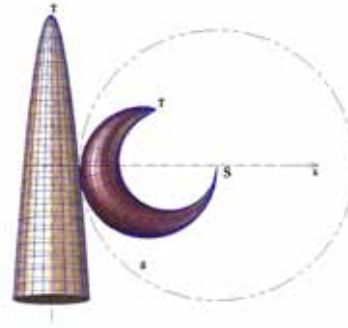


Figure 13: Surface Model of the revolving paraboloid and the crescent-shaped surface of the 4th order

5. DIAGRID STRUCTURAL SYSTEM

Diagrid is a structural system which consists of

triangular structures with diagonal support beams. This system has been widely used for newly-constructed high-rise steel buildings because of its exceptional structural efficiency and aesthetic potential. There is no structural system superior to diagrid concerning the efficiency in the distribution of load among all the members of a structure [11]. This structural innovation allows the construction of tall buildings in forms that we previously thought impossible. It uses a number of units that are combined to make a grid with a myriad of different structural possibilities and to maximize the efficiency of the most commonly used steel units. The use of this system requires less structural steel, thereby reducing the weight and cost of buildings, and at the same time allows the transfer of significant vertical load. However, the grid itself is not strong enough to carry the horizontal load, so a ring system has to be added to prevent bending. Diagrid can be used for a variety of geometric shapes and it has enormous potential for the future architecture. This can be seen in Figure 4, which was taken from the site: <http://plusmood.com/2008/11/capital-gate-rmjm-architects/>. This project used Tekle-structured software which has been used for modeling some of the world's tallest and most popular buildings. This software uses three-dimensional modeling. It calculates three-dimensional coordinates for each component of the diagrid grid, which ensures a flawless integration and assembly of the main frame and the supporting beams.



Figure 14: Capital Gate of the architectural company RMJM in Abu Dhabi

Diagonal structures were invented by Russian engineer Vladimir Shukhov. He was famous for his pioneering work on new analytical methods in different areas. Shukhov left a permanent mark on early constructivism in Soviet Russia. As a leading engineer and mathematician in the late 18th and early 19th century, he created world's first hyperboloids as well as the thin shell of tensile structures of exquisite sophistication and elegance.

In his paper Chao Huang [4] examines the carrying capacity of the diagrid structure depending on the angle between the diagonals and it varies up to 30%.

There is an increasing number of high-rise buildings around the world [1] that dominate the landscape. Their construction is possible due to the invention of the Diagrid system and rapid development of construction equipment, mechanization and materials [10]. Future research studies should investigate structural systems that are efficient in seismic zones. Advantages of the Diagrid system compared to conventional systems lie in its large load-bearing capacity and in the structure of the shell that allows large openings on the buildings and uses 20% less steel than conventional steel structures. It can be concluded that the diagrid system would be particularly suitable for the surfaces obtained by inversion, because they have two systems of circular cross-sections and the parts of these surfaces can be used to make different shapes and free-forms.

6. CONCLUSIONS

With different appointments of the sphere-s and the center of harmonic symmetry S, we can transform surfaces of lower orders into surfaces of higher order with attractive forms that are suitable for constructions because of their systems of circular cross-sections obtained by mapping. To construct buildings modeled in this way, in a system such as Diagrid, it is necessary to find optimal directions of isolines, which would represent the directions of the supporting beams of the diagrid structural sys-

tem and optimize its capacity. This should be the direction of future investigations.

Apart from clearly defined geometric surfaces, free forms also play an increasingly important role in modern architecture. These forms are most commonly obtained by using the geometry of NURBS. However, the construction of these structures presents a great problem.

Recent technological advances made it possible to mass-produce doubly curved panel boards that can wrap architectural free-form surfaces.

M. Eigensatz et al [12] created a computer algorithm for the production of large-scale architectural freeform surfaces obtained using NURBS geometry. The entire surface of a building is divided into panels, whose shape and size depend on the shape of the area to be wrapped. The shape and smoothness of the surface to be wrapped must not be distorted. Such panels only minimally alter the projected surface, and reduce the cost of building.

Freeform glass structures are commonly used in modern architecture. The area is divided into cylindrical (84%), flat (10%), and double curved (only 6%) panels, which means that 94% of the panels can be produced using an automated industrial process and retain the curvature of the surface designed by an architect [2].

ACKNOWLEDGMENTS

The research is financially supported by Ministry of Science and Education, Republic of Serbia, under the project TR-37002, New bioecological materials for protection of soil and water.

REFERENCES

- [1] Ali M. M. and Moon K., "Structural Developments in Tall Buildings: Currents Trends and Future Prospects", *Architectural Science Review*, 50.3, 2007, pp 205-223.
- [2] A. Schiftner, M. Eigensatz, M. Kilian, Large scale double curved glass facades made feasible-*The Arena Corinthians West Facade July 1st, 2013* (Glass Performance Days 2013)
- [3] B. Benton, Advanced Auto-CAD 2011, *Infinite Skills*, 2011
- [4] Chao Huang, Xiao-lei Han, Jing Ji, Jia-min Tang (2010): Behavior of concrete-filled-steel tubular planar intersecting connections under axial compression, Part 1: *Experimental study, Engineering Structures* 32 (60-68).
- [5] G. Đukanović, J. Maksić, Pencil of Conics as Different Plane Intersections of Pencil of Quadrics, *Proceedings of 10th ICGG, 28 July - 2 August 2002, Kiev, Ukraine, 2002*, pp 177-181
- [6] G. Đukanović, V. Matić, Graphic transformation of hyperbolic pencils of circles into pencils of conics and these into pencils of curves of the 3rd or 4th order, *Proceedings of 2nd international Scientific Conference moNGeometrija 2010, Beograd*, pp 81-89.
- [7] G. Đukanović, M. Janić, Graphic transformation of "asymmetrical" pencils of conics into pencils of curves of the 4th and 3rd order, *Proceedings of 3rd International Scientific Conference moNGeometrija 2012., Novi Sad*, pp 353-360
- [8] G. Đukanović, M. Obradović, The Pencil of the 4th and 3rd Order Surfaces Obtained as a Harmonic Equivalent of the Pencil of Quadrics through a 4th Order Space curve of the 1st Category, *Facta Universitatis, Series: Architecture and Civil Engineering*, VOL.10 No 2.2012, pp. 193-207.
- [9] G. Đukanović, The pencils of curves of the third and fourth order obtained by mapping the pencils of conics, *Doctoral dissertation, Faculty of Architecture, Belgrade*, 2012
- [10] K. Moon, J. Connor and J. E., Fernandez "Diagrid Structural Systems for Tall Buildings: Characteristics and Methodology for Preliminary Design", *The Structural Design of Tall and Special Buildings*, Vol. 16.2, 2007, pp 205-230.

- [11] K. Moon, "Design and Construction of Steel Diagrid Structures", *School of Architecture, Yale University, New Haven, USA*, 2009.
 - [12] M. Eigensatz, M. Deuss, A. Schiftner, M. Kilian, N. J. Mitra, H. Pottmann, Mark P. Case Studies in Cost-Optimized Paneling of Architectural Freeform Surfaces
 - [13] M. Stavric, Harmonic Synthesis and Constructive Processing of the Surfaces of the Higher Order, *Doctoral dissertation, Faculty of Architecture, Belgrade*, 2002
 - [14] L. Dovniković The Harmony of the Spheres, The Relativistic Geometry of Harmonic Equivalents, *Matica srpska, Novi Sad*, 1999
 - [15] L. Dovniković, The Relativistic Geometry of Harmonic Equivalents, *Proc. 6th ICECGDG, Tokyo*, 1994, pp305-309.
 - [16] L. Dovniković, Relativistic Homology as a Way of Tying or Untying Singular Points, *Journal for Geometry and Graphics, Vienna*, Volume 8, No.2, 2004, pp 151-162
 - [17] M. Obradović M., B. Popkonstantinović, S. Mišić and M. Petrović, "Possibilities of Deltahedral Concave Cupola Form Application in Architecture", ICEGD 2011- "Sustainable Eco Design", IASI, Romania, *The 4th International Conference on Engineering Graphics and Design*.
 - [18] S. Krasić., "Geometrijske površi u arhitekturi", monografija, Niš, Galaksija, Građevinsko – Arhitektonski fakultet Univerzitet u Nišu, 2012.
 - [19] Shambina S., Sazonov K.: Application of analytic surfaces and bionic forms in architectural design, *Proceedings of 3rd International Scientific Conference moNGeometrija 2012.*, Novi Sad, 2012, pp 499-506.
- of Forestry. She has graduated in Civil Engineering at the University of Belgrade. In her scientific research, she is dealing with descriptive and relativistic geometry. She has created possibilities for the mapping curves and surfaces and obtaining new forms which can be of great use both in the theory of geometry and in the practice of architecture.
2. Đorđe Đorđević PhD, Assistant professor at the University of Belgrade – Faculty of Architecture, teaching Geometry of architectural form. He has graduated in Architecture at the University of Belgrade. In his research, he is dealing with two scientific fields: Constructive Architectural Geometry and Architectural Close-range Photogrammetry.
 3. Marija Đ. Obradović PhD, Associate professor at the University of Belgrade - Faculty Civil Engineering, teaching Descriptive Geometry and Computational Geometry. She has graduated in Architecture at the University of Belgrade. In her scientific research, she is dealing with various fields of geometry, especially with Polyhedral Structures in Architecture and Engineering.
 4. Slobodan Mišić PhD, Assistant professor at the University of Belgrade - Faculty Civil Engineering. He has graduated in Architecture at the University of Belgrade. In his scientific research, he is dealing mostly in General Collinear Planes, Constructive Geometry and Polyhedral Structures in Architecture and Engineering.

ABOUT THE AUTHORS

1. Gordana Đukanović PhD, Assistant professor at the University of Belgrade – Faculty

APPLICATION OF THE CROSS-RATIO TO THE ANALYSIS OF ARCHITECTURE

Riccardo MIGLIARI, Leonardo BAGLIONI
Sapienza University of Rome, Italy

ABSTRACT: Rudolf Arnheim, in his essay dedicated to the dynamics of architectural form, describes in a simple and clear way the two possible perceptions of an architectural space: *as it is* and *as it looks*. The architecture, in fact, appears to our eyes seconding the laws of perspective and appears to our mind in its solid formal regularity. Actually, between these two opposite poles exists a means, which allows to pass with continuity from one to another and which, as we will see, moves towards one another in a metamorphosis without continuity solution. When we observe a perspective that simulates the depth of an architectural space or that simply alludes to that depth, inevitably is activated the visual perception and the perception of mental space and the dialogue between them develops the continuous transformation and engages, therefore, at the same time, the idea of movement. The critique of architecture has recently emphasized some of the façades of Roman buildings, and not only, which, using the above relations, induce, in the onlooker, illusory impression of over-hangings and depths. Sandro Benedetti called these “pulsating façades”, referring to the sensation of motion that they evoke. The historical collocation of these buildings is the same as the discovery of the perspective, of its first codification and its refinement, and it therefore seems right to employ analysis of projective type to verify whether the perspective could have given rise to these compositional solutions. The study that we are now presenting began with the measurement of the intervals (“*as they are*”), apparently casual, that line these façades, in order to find the perspective rule that generated them and to define the depths that said intervals are capable to evoke (“*as they seem*”), depending on the observer's position. Relations between real and virtual distances, can be implemented in a simple instrument (using a graphical algorithm editor) which can be used to analyze the “virtual” façade (created by the perception of an observer who changes his position) of Roman building. A more detailed analysis of the cross-ratio of the above-said intervals enables instead the measurement of the acceleration of the perspective view, or better, their variation, which is in fact the characteristic that produces the sensation of a “pulsation”. This method is based on the fundamental characteristic of cross-ratio to be an invariant for projective transformation. Since the cross-ratio for four points which separate three segments with equal intervals, measure $4/3$, we can analyze how much the cross-ratio of adjacent segments diverge from this value to quantify the pulsation of the façade. The comparison between obtained results with the two analysis methods, the pictures of the façades taken from different points of view and measured elevations demonstrate the existence of the perspective matrix in the façade composition.

Keywords: Perspective, Cross-ratio applications, Pulsating façade, Visual perception.

1. PERCEPTION OF ARCHITECTURE

Rudolf Arnheim, in his essay dedicated to the dynamics of architectural form [1], describes in simple and clear way two possible

perceptions of an architectural space: *as it looks* and *as it is*. The architecture, in fact, appears to our eyes seconding the laws of perspective but appears to our mind in its solid formal regularity. Who enters a Renaissance and Baroque

church has no doubt that the pillars that support the broad arches of the nave are all of equal height, but he can not help observing the apparent convergence of the frames in the altar, as well as any other scenic effect, and he can not help being emotionally involved .

The issue was already known and debated in the Renaissance, when it was encoded the “legitimate construction” and, simultaneously , the first theoretical discussions about the representation of architecture. The circle of artists who hides behind a letter to Leone X was well aware of the ambiguity mentioned by Arnheim: if they had to survey and draw the Roman survivor antiquities, they refused the way of pictorial perspective which shows things as they appear, to the benefit of plan, section and elevation, which show for what they are.

Apparently, the perception of architecture moves between two extremes: “*as it looks and as it is*”, to quote Arnheim. But, actually, between these two opposite poles exists a vehicle, which allows to pass with continuity from one to another and that, as we shall see, moves one towards the other in a metamorphosis without continuity solution. To realize this vehicle immediately, just think of the apse of Santa Maria in San Satiro: here the distinction between as it looks and as it is, is no longer so simple and clear. In fact, the “*as it looks*” alludes to an “*as it is*” wholly illusory. The real object, in this case, acts as intermediary between the visual perception (perspective), and the mental perception (orthographic), which, in turn, is responsible of the illusion of depth, in the continuation of the nave beyond the transept. According to the studies of the *Transactionalist School of psychology* [2], in fact, is the comparison between the known models of a stereo-metric space and the perspective image, which induces in the mind the illusion of depth . Who knows these features of perception is thus able to control perfectly the illusion, as demonstrated by Ames’s experiments.

Let’s get back to the idea of the movement. The apse by Bramante is, in fact, a phase of the movement that transforms the *as it is* in *as it*

looks, that is the church choir in the vision we have of it from main nave. The vehicle, or solid perspective, made in the slim thickness of 97 cm, it is just a phase of transformation, as only one frame in the sequence of a film. But the transformation is continuous and has as its extremes: on the one hand , the space of regular shapes we call orthographic, which we recognize in the plan and elevation drawings of the letter to Leone X, and on the other the flat perspective which is obtained by canceling three-dimensional perspective space in the two dimensions of a drawing.

At a closer look, every flat perspective, contains this potential. In fact, in one perspective image, without its geometric code, correspond infinite three-dimensional realities. And the transformation that translates ones into others infinite possibilities, it is still continuous and can be seen as an action film. When we place ourselves in a perspective that simulates the depth of an architectural space or simply alludes to that depth, so inevitably, it activates the visual perception and the perception of mental space and in the dialogue among them, the continuous transformation which we talked about, develops and engages, therefore at the same time, the idea of movement.

This is what we are going to demonstrate with a few examples.

It is not uncommon the case of historical buildings where the façade is not composed according to a regular scan of the space, but according to a non uniform scan, which seems to allude to perspective intervals and that, in fact, is able to conjure up illusions we mentioned and, in particular, the idea of movement. Quite appropriately, Sandro Benedetti has decided to define these façades, *pulsating façade* [3] [4]. Is it possible to analyze these simple allusions to the perspective, in order to describe geometrically the feelings that are able to evoke?

To answer this question, we should start from a consideration both simple and obvious: the regular scan of a façade (Fig. 1) (with the axes of the openings arranged in equal intervals)

turns, when it is observed in perspective, in a perspective scan, in which the same intervals follow the law of *reduction of apparent magnitudes* (Fig. 2).



Figure 1: regular scan of a façade



Figure 2: reduction of apparent magnitudes

Reason why, if you want to create the illusion of movement of the front, in the sense of spatial depth that moves away from the observer, will be sufficient to have the axes of the

openings (or other equally strong compositional element), at declined intervals, rather than regular.

But when, and how, it happens that a series of irregular intervals can be said to be properly declined in perspective illusion?

The question is not trivial, because if there is a unintuitive matter in perspective, this is precisely the reduction of apparent magnitudes. The convergence of the images of parallel lines to a point is a phenomenon that can easily be seen (just using a mirror as the ancients knew) and you can easily play with (just using a nail on the part to “break”). But if we try to do the same with the gradual reduction of the images of objectively equal intervals, known as reduction, then the common sense is not enough and you need the geometric science. In fact, even if you generate the reduction thanks to the mirror, the law that governs it is not at all obvious, as it is instead, a clear convergence of intersecting lines at the same point. And any intuitively attempt to reproduce that reduction in the drawing fails, because it conflicts with the corrections induced from physical experience and its perception of mind, which tell us that the intervals are equal, while the visual experience, at the same time, tells us that these intervals are reduced. The result of an intuitive reduction, such as the one that tries an unconscious drawer, is, in fact, a succession of intervals that exceed or do not sufficiently contract.

So, if you want to look for a test of knowledge of the ancient perspective, it would be simplistic to look only in the presence of a single vanishing point, as Panofsky did on the frescoes of the second style [5]. Instead, it is much more useful and convincing investigate the reduction of apparent magnitudes to check if it is correct or it is not according to the laws of perspective [6].

That said, we want to provide some useful tools for analysis of the reduction and therefore also for an analysis of the movements induced by tricks of perspective on the façades we dealt with.

We will use two method: the first one based

on a simple direct comparison between intervals and the second one based on an examination of the cross-ratio values that generate the same intervals.

2. FIRST METHOD: DEPTH ANALYSIS

Imagine dividing the façade into as much segments as there are intervals between the windows axis. These segments will have the rectangular aspect of as many wings of a theatre stage, all located in the same plane. If all intervals are equal, the corresponding segments of the façade are perceived in the position where they are, and that is like wings belonging to the same plane. If, however, one segment is smaller than the largest, it is perceived as if the segment façade was further away, while the largest interval “stays” on the façade.

The variation of the (virtual) distance of a shorter interval can be calculated as follows: set the focal length f (or principal distance) equal to the unit (which we can assume equal to the distance of the observer from the front), it is shown that one object distant $d=f$ from observer appears in true form, one object at $d = 2f$ is reduced by half, one object at $d=3f$ is reduced by

one third, and so on, according to the relation:

$$a' = \left(\frac{a}{d} \right) \cdot f \quad (1)$$

where a' is the width of reduced interval, a is the width of the reference integer interval (the largest among those present), d is the virtual distance (distance of the illusory plane which hosts a' from the parallel plane passing through the observer) and f is the distance of the observer from the façade, which serves as module or unit of measurement. In fact (Fig. 3), if O is the observer and AB is a segment located on the projection plane at distance f from O , a segment PQ (equal to AB but far from O at $d=2f$) is projected on the plane in a segment $P'Q'$ which is half of the size of PQ , as can be seen by comparing the two similar triangles OQP and $OQ'P'$ in which OQ is double of OQ' for construction and, consequently, also the sides PQ and $P'Q'$ are in the same ratio.

A similar reasoning allows to establish that $R'S'$ is one third of RS which is distant $d=3f$ from O , and so on. Therefore, the (virtual) distance d of any illusory segment a' , painted or drawn on projection plane (the façade plan) can be calculated with the expression:

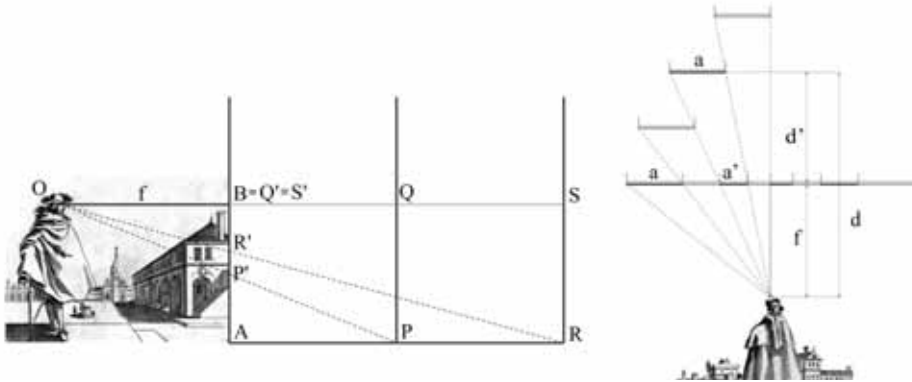


Figure 3: variations of virtual distances

$$d = \left(\frac{a \cdot f}{a'} \right) \quad (2)$$

while the distance d' of the virtual wing a' from the façade, measures evidently $d \cdot f$ and therefore it is obtained by:

$$d' = \left(\frac{a \cdot f}{a'} \right) - f \quad (3)$$

Now imagine a designer who, well aware of the laws of perspective, has desire to give the façade of the building which is planning an effect of concavity .

Divided the front in wings of equal width a , b , c , d , e (in grey), he moves back b , c and d as appropriate (Fig. 4).

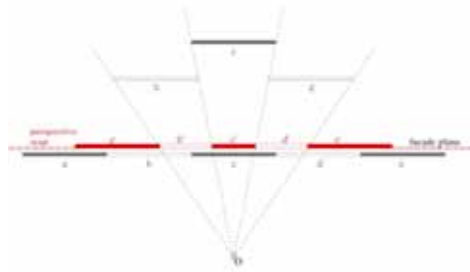


Figure 4: creation process of concave façade

At this point, choosing a central location of the observer O , he projects backward on the plane façade the ends of the wings, getting the segments b' , c' , and d' , with perspective reduction. With these segments and with the extremes a and e (which have remained in true measure) our designer can build a scale (that is a scan of the façade with new modules a' b' c' d' e') that will serve as a model for the composition of the façades. Applying this scale to the note construction that divides a segment into equal or proportional parts, the designer will get a scheme of a façade divided into perspective parts. The process, however, remains embedded in the composition result so that we (viewers of finished result) can easily rebuild it using the formula (3) to the intervals in which

the façade is divided.

The formula returns the illusory distance of the central segments to the façade and, with them, a table that illustrates in a symbolic but expressive way, the concave trend of virtual façade (Fig. 5). Relations between real and virtual distances, can be implemented in a simple instrument (using a graphical algorithm editor) which can be used to analyze the “virtual” façade (created by the perception of an observer who changes his position) of a building.

Taking advantage of this device, we can also verify a peculiar feature of this optical correction, as well as perspective in

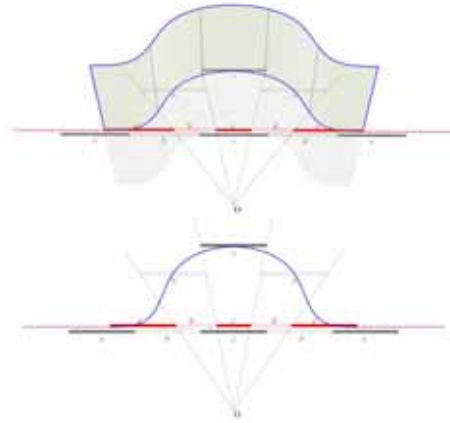


Figure 5: virtual façade of a building

general: the intensity of the effect depends on the observer's position. In fact, the effect of concavity is stronger the more the observer is far from the front and decreases with distance to zero when the observer crosses the threshold and enters, so to speak, the illusion (Fig. 6).

It's time to move from theory to practice, examining a building that presents the characteristics of the pulsating façade and precisely the *Lateran Palace* in the elevation overlooking Piazza di Porta San Giovanni in Rome, designed by Domenico Fontana. In this pulsating façade, the largest of the intervals f measures 8.10m and therefore the above formula gives

the following results :

Table 1.

Interval	Width (m)	Virtual Distance from façade plane (m)
<i>a'</i>	6,57	2,328767
<i>b'</i>	6,63	2,217195
<i>c'</i>	6,93	1,688312
<i>d'</i>	7,95	0,188679
<i>e</i>	8,1	0
<i>f'</i>	7,88	0,279188
<i>g'</i>	7,6	0,657895
<i>h'</i>	6,86	1,80758
<i>i'</i>	6,68	2,125749
<i>l'</i>	6,83	1,859444

The table reads as follows:

the first column gives the name of progressive intervals, from left to right; the interval *f*, being the largest, was hired as belonging to the plane of the façade;

the second column gives the width of the intervals expressed in meters;

in the third column figures the result of the formula (3), which expresses the distance from the façade to which each segment is perceived with respect to the reference, which has null distance.

This means, for example, that an observer placed in front of the façade at 10 meters, perceives the segment *a'* as if it were about 2,32 meters farther away. Of course, this value increases if the observer moves away and decreases if the observer approaches, as has already been said. It should be noted, however, that the greater distance also means, of course, a “crushing” perspective, as it happens, for instance, in photographs taken with a telephoto lens. And so the two opposing effects balance each other, making the choice of the distance of the observer from the façade, at the end of the calculations that we have illustrated, merely conventional.

If we put the above values in a graph, in which the *X-axis* is aligned with the façade and

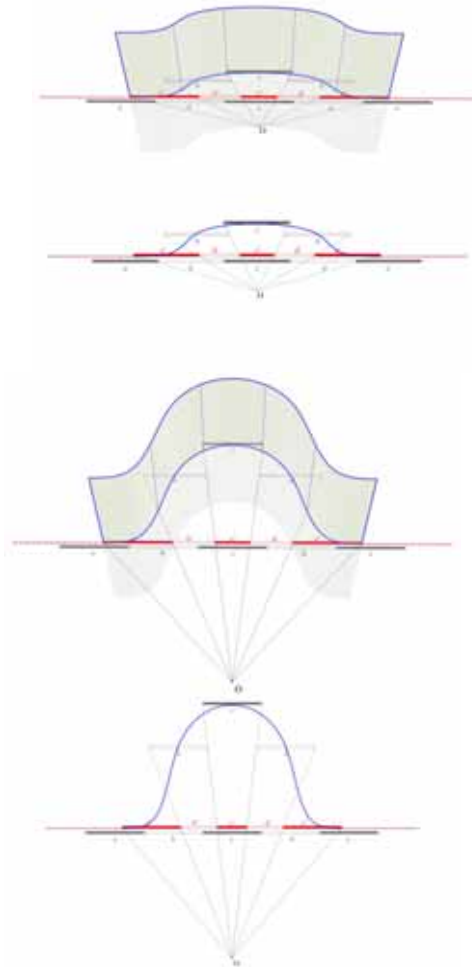


Figure 6: variations of intensity of concavity depending on the observer's position.

Y-axis indicates the illusory depth, we obtain a drawing that describes significantly the succession of concavity and convexity produced by the perspective correction (Fig. 7). The continuity of the masonry then merges together these illusory depressions and recompose impression of motion that has been called “pulse”. You can read even more about this effect by interpolating the data with a curve, taken back to the front of the building.

3. SECOND METHOD: CROSS-RATIO INVARIANCE

The eastern façade of the Lateran Palace, demonstrates how the application of perspective corrections we have illustrated is more complex than one might expect. In fact, the

continuous variation of the intervals is used not just to create an effect of concavity or convexity, but to switch to one another, with greater or lesser intensity.

It is opportune to think of a more appropriate method of analysis that provides a measure of the variation of the effect of perspective. This method is based on the cross-ratio invariance. To understand how it works, we must remember that if a series of contiguous segments are aligned and subjected to an operation of central projection (as it happens, in fact, in perspective and in human vision), their measures and their mutual relations are as a result of the projection. For example, if you build the perspective of three segments **AB**, **BC** and **CD**, aligned and contiguous and of equal length (Fig. 8), their lengths result altered and, generally, also the relationships you can identify by comparing a segment with each other. However, there is a special relationship, which does not suffer any alteration: it is called *cross-ratio* and

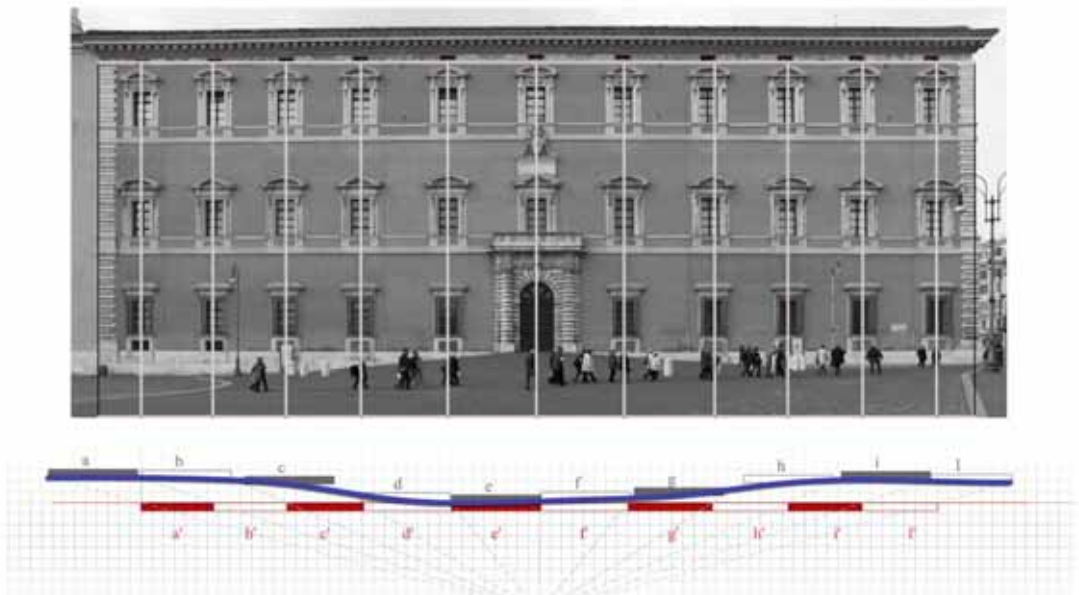


Figure 7: pulse analysis of eastern façade of Lateran Palace

it is an invariant in projective transformations.

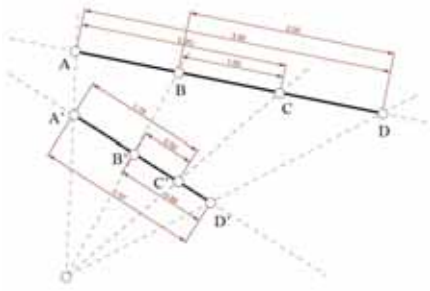


Figure 8: cross-ratio in equal segments

We see, then, how to build and measure the cross-ratio. Consider extremes **A**, **B**, **C** and **D** of the three segments aligned and contiguous that we have described above, without excluding, however, that may be different from each other and of any length.

We can build the following simple ratios:

AC/BC and **AD/BD**

The cross ratio (briefly indicated with **(ABCD)**) is the ratio of simple ratios constructed as above, that is:

$$\frac{\frac{AC}{BC}}{\frac{AD}{BD}}$$

Suppose now that the three intervals **AB**, **BC** and **CD** are all equals to the unit and ask what will be, in this case, the value of the cross ratio. In the event that we hypothesized the segment **AB** measures $2 \cdot (\mathbf{AB} + \mathbf{BC})$, the segment **BC** is worth the unit, and the segment **AD** is worth $2 \cdot (\mathbf{AB} + \mathbf{BC} + \mathbf{CD})$, the segment **BD** is still worth $2 \cdot (\mathbf{BC} + \mathbf{BD})$. Therefore, the cross-ratio is:

$$\frac{\frac{AC}{BC}}{\frac{AD}{BD}} = \frac{2}{3} = 2 \cdot \left(\frac{2}{3} \right) = \frac{4}{3}$$

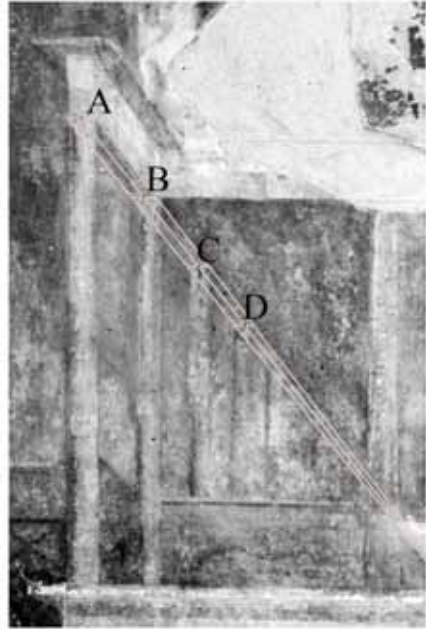


Figure 9: painting in House of Augustus

Thus, the cross ratio of four points separating equal intervals worth $4/3$, that is $1,33 \dots$ But since the cross-ratio is an invariant in projective operations, even any perspective of these three segments aligned and contiguous, will provide the same result.

Consequently, the knowledge of this value allows an immediate verification of the correctness of a perspective: if any of the three intervals that in real world represent three equal intervals between them, have a cross-ratio equal to $4/3$, the construction of the reduction of apparent magnitudes is legitimate. It respects exactly the theoretical principles of central projection. Note that this is always, whatever is the point of view, and how many times those three segments aligned and contiguous have been subjected to projection operations. If, for example, we measure the cross-ratio of four points on one of the paintings in the House of Augustus on the Palatine, we can verify that this value is very close to the theoretical value

of 1.33... (Fig. 9). This observation could perhaps lead us to revise the reading given by Panofsky about perspective of the ancients. If we repeat the measurement on any photograph of the fresco, also taken from a lateral and oblique point of view, the value of the cross-ratio does not change, precisely because it is invariant in projection operations.

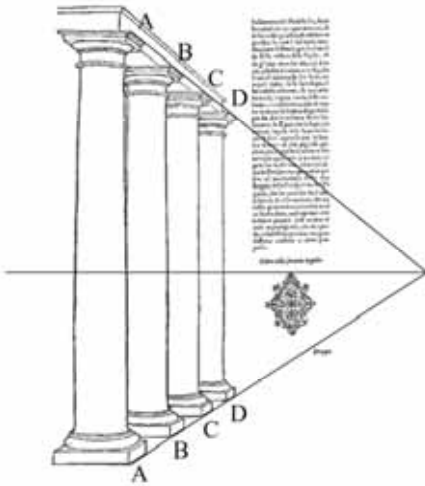


Figure 10: Vignola's drawing of four Tuscan columns

The same check conducted on the drawing of four Tuscan columns, which is located on p. 142 of the edition of the Two Rules of Vignola in 1583[7], provides similar results (Fig. 10).

But these are just simple examples of the use of the cross-ratio as a tool for qualitative analysis of perspective.

The case that we are examining is more complex, because here the searched effect is not that of a constancy in succession in the depth of the intervals, but of a variation such as

to bring forward what was just before backward and vice versa. We are therefore in a situation that sees varying the widths of the illusory intervals created by perspective and, therefore, sees increasing and decreasing the value of the cross-ratio compared to the value of $4/3$ that describes the constancy of the intervals.

This change introduces a new concept, which we call *perspective acceleration*.

Consider, therefore, some wings of equal width that, instead of being arranged all on a plane, are arranged in space so as to describe a sinuous shape, as seen in the figure (Fig. 11), which for simplicity is in top view.

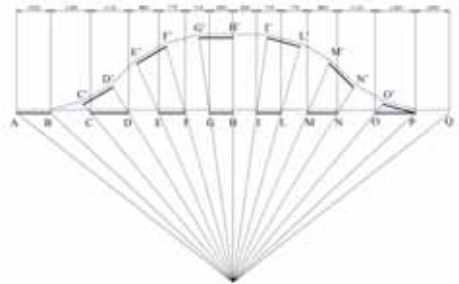


Figure 11: design process of a pulsating façade

We may use this model as polygonal chain of the movement that we want to impose to a façade. Imagine, then, that the façade is represented by the straight line **AQ** and project, from a center of projection **S**, the vertices of the polygonal chain on the straight line **AQ** at points **A, B, C, ... O, P, Q**. We will obtain, in this way, the widths of eleven intervals that punctuate the façade according to a perspective rhythm. Applying to this configuration the method that we have proposed for the analysis of illusory depth, we get the graph shown in Figure 12 (Fig. 12).



Figure 12: illusory depth analysis

As you can see this chart proposes the sinuous course of illusion, tracing the starting model, except in the extreme. These appear almost retreated to accentuate the effect that we wanted. This happens because the projections of related segments are shorter than the projections of the segments that follow (on the left side) and precede (to the right side) and, therefore, appear as retreated respect to the plane of the façade. This effect tends to disappear, increasing the observer's distance.

We now want to consider the acceleration of the perspective illusion. For this purpose, consider the fourteen segments of the polygonal chain, and its fifteen vertices **A, B, C, D, E, F, G, H, I, L, M, N, O, P, Q**.

With these points is possible to build twelve quadruples:

ABCD, BCDE, CDEF, DEFG, EFGH, FGHI, GHIL, HILM, ILMN, LMNO, MNOP, NOPQ. Each of this quadruple includes three intervals of unequal lengths, which give rise to the cross-ratios: **(ABCD), (BCDE), ... , (MNOP), (NOPQ)**.

We would like to mention that these cross-ratios are equal to those of the angles subtended by the points that generate them in the viewer's eye. Reporting on the *X-axis* the natural succession of intervals and on the *Y-axis* the values of the relative cross-ratios, you get, based on the principles we have set out, a graph that describes the acceleration of perspective illusion applied to the façade (Fig. 13).

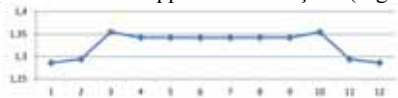


Figure 13: perspective acceleration analysis

This chart does not describe the shape of the virtual façade, but the intensity of its “pulse”. Quick changes in the cross-ratio signify a sudden accentuation of depth, as is done in extreme ranges, while delicate variations describe a regular pattern, as is the center of the façade.

To apply these theoretical considerations to a particular case, we will examine again the

eastern façade of the Lateran Palace. Here, however, the number of intervals is odd, because, as almost always happens, the central one is dedicated to the portal of entry. In cases like this, you can not build many quadruples of points to cover the entire length of the façade, because the range will always be excluded. It is possible, however, to build two series of quadruples proceeding from left to right and from right to left. The central quadruple, while having the opposite trend, for example **(DEFG)** and **(GFED)**, will have the same value, because this is a known property of cross-ratio [8].

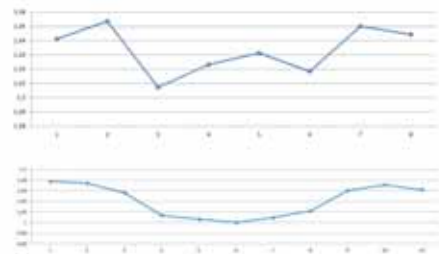


Figure 14: depth analysis and perspective acceleration of eastern façade of Lateran Palace

This method, applied to the façade of the Lateran Palace, gives this result (Fig. 14) which here is compared with the previous depth analysis. In the figure the lower graph describes the deformation produced by the illusory perspective of the façade, while the upper describes, interval after interval, the intensity of this changes, that is the vibration of the façade or its “pulse”, to use Sandro Benedetti’s expression.

4. CONCLUSIONS AND FUTURE WORK

The presented investigations, show the evidence of perspective matrix in the composition of the façade, they can measure the intensity, they demonstrate the trend, but nothing they say yet about the theoretical tools, reasoning and methods used to obtain these qualities.

The legitimate questions that stem from this

gap will have to answer a further study which, based on a careful metrology analysis, could explore the relationship between integer measures, as the graphic constructions that can establish such relationships.

For example, the construction of reduction of apparent magnitudes based on “point of distance”, described by Vignola in the Two Rules, can lead to relationships between integers such as those expressed by sets of three segments:

5, 3, 2, 10, 5, 3, 10, 6, 4, 20, 12, 8, 24, 18, 14, 30, 18, 12, 36, 27, 21 ... etc. and multiples, where all are able of 4/3 cross-ratio. In fact given two segments aligned and contiguous, it is possible to build a third, a fourth and nth segment so as to form a perspective sequence with constant speed, that is to form the perspective of equal segments in real space. Given three intervals aligned and contiguous, **AB**, **BC** and **CD**, the value of the third interval is obtained by solving, compared to **CD**, the equation:

$$\frac{\frac{AC}{BC}}{\frac{AD}{BD}} = \frac{4}{3}$$

Where:

$$AC = AB + BC$$

$$AD = AB + BC + CD$$

$$BD = BC + CD$$

for which:

$$\frac{\frac{AB + BC}{BC}}{\frac{AB + BC + CD}{BC + CD}} = \frac{4}{3}$$

that is:

$$CD = \frac{(AB \cdot BC + BC)^2}{(3AB - BC)}$$

So it can not be excluded that beginning with a regular scan of “constant speed”, Domenico Fontana has also introduced some simple corrective to avoid fractional measures and dosing the illusory movement of the façade according to the sensitivity of his eye and the mastery of his art.

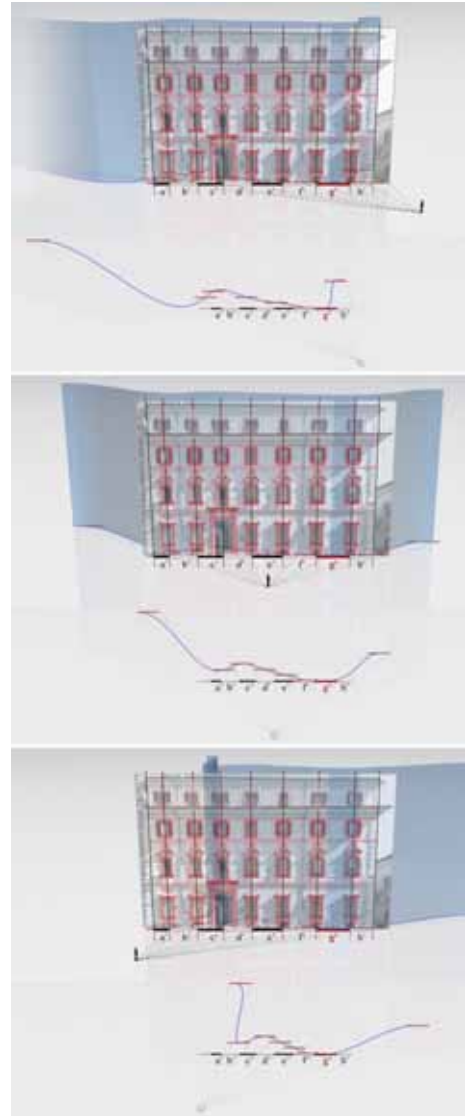


Figure 15: pulsating façade perception in Capizucchi Palace (*Giacomo della Porta, 1593*)

ACKNOWLEDGMENTS

Authors would like to acknowledge Michele Calvano who contributed to the development of parameterization algorithm used in this research.

REFERENCES

- [1.] R. Arnheim, *The Dynamics of architectural form*, University of California Press, Berkley - Los Angeles - London, 1977.
- [2.] F. P. Kilpatrick, *La Psicologia transazionale*, Milano, Bompiani 1967.
- [3.] S. Benedetti, *I palazzi romani di Giacomo della Porta*, in *Roma e lo studium urbis: spazio urbano e cultura dal Quattro al Seicento*, edited by Paolo Cherubini, 441-70, Rome, 1992.
- [4.] B. Azzaro, *Facciate pulsanti e spazio urbano nella Roma del Cinquecento*, in *Quaderni dell'Istituto di Storia dell'Architettura, Dipartimento di Storia, Disegno e Restauro dell'Architettura*, Nuova Serie, F. Cantatore, et al., Editors. 2013, Bonsignori Editore: Roma. p. 209-222.
- [5.] E. Panofsky, *La Prospettiva come "forma simbolica" e altri scritti*, 1966, Milano.
- [6.] R. Migliari, *Panofsky and Perspective*, in *Disegnare, idee immagini, ideas images*, vol 31, pp 28-43, Roma, Gangemi Editore, 2005.
- [7.] I. D. Vignola, *Le due regole della prospettiva pratica*, Casso di Risparmio di Vignola, 1583.
- [8.] O. Chisini, *Lezioni di Geometria Analitica e Proiettiva*, Zanichelli Bologna, 1967, I, 9.

ABOUT THE AUTHORS

1. Riccardo Migliari is Full Professor in *Fundamentals and Applications of Descriptive Geometry* since 1990 at the Faculty of Architecture at the 'La Sapienza' University in Rome. He is assiduously engaged in research, particularly in the areas of *Descriptive Geometry* and

of *Representation and Instrumental Survey of Architecture*. He directed, as Scientific Manager, the architectural survey of Coliseum in Rome during the preliminary studies for the restoration of the monument undertaken by the Archaeological Superintendence of Rome. From year 2003 he deals, in particular, with the renewal of the studies on the scientific representation of the space, within in the evolutionary picture of the descriptive geometry, from the projective theory to the digital theory and from the graphical applications to the digital modeling.

2. Leonardo Baglioni, is Assistant Professor in *Fundamentals and Applications of Descriptive Geometry* at the Faculty of Architecture at the 'La Sapienza' University in Rome. His research focuses on issues related to descriptive geometry and renewal through the introduction of methods of digital representation. In this field of research, he is interested in new applications of descriptive geometry with reference to the development of the recent achievements of discrete differential geometry which today play a significant role in contemporary architectural projects.

AN APPROXIMATION METHOD OF TRIANGULAR MESH MODELS FOR GENERATING DEVELOPMENT DIAGRAMS WITH CREASES AND SLITS

Kuniko Takahashi¹, Tomohiro Tachi² and Yasushi Yamaguchi³

¹The University of Tokyo, Japan ²The University of Tokyo / JST PRESTO, Japan

³The University of Tokyo / JST CRESTO, Japan

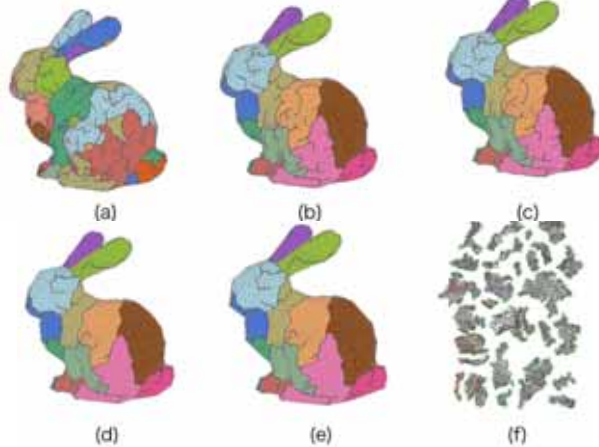


Figure. 1: The Steps of Our Method.

ABSTRACT: Making paper crafts from three-dimensional mesh models is useful not only in entertainments but also in architecture or in other engineering applications. There have been many studies how to approximate the mesh model under a constraint of paper's property, known as developability. In those studies, the shape of each part is aimed to be a circular disk in order to maintain its manufacturability. In this paper, we propose a method to approximate a triangular mesh model with as little number of the parts as possible. The method achieves high accuracy with a few parts by allowing creases and slits within parts.

Keywords: papercraft, development, slits and creases.

1. INTRODUCTION

It becomes popular to process or display three-dimensional shapes with 3D computer graphics techniques. Three-dimensional mesh model is one of the most popular representations to describe 3D shapes. It requires a huge amount of small triangles to represent a smooth surface.

Papercraft is one of the ways to easily obtain a physical model. Many people enjoy creating papercrafts as their hobby. It is useful to evaluate the shape with a physical model but also to realize a final product as an architectural structure or a machinery part by using metal sheets instead of papers. However, the shapes of papercrafts are limited to developable shapes due

to the nature of papers. It is characterized that papers have no stretch. There have been studies how to approximate the mesh model under this constraint of paper's property, known as developability.

When we reproduce a mesh model as a papercraft, it is necessary to develop the model on to a plane, namely, its development. Various methods to make this development have been proposed. Most of them adopt the approach that a mesh model is separated into small areas, called charts, then each chart is approximated by a developable surface. It is required that the target shape is well approximated by the charts and is realized in an easy-to-glue manner. The previous methods define easiness of gluing as a number of charts and simplicity of chart boundaries like nearly circle or having smooth edge.

In this paper, we propose an approximation method that is aimed at reducing the number of charts by introducing slits and creases into the charts. To achieve the method, we use the angle constraint around vertices so that a total angle around a developable vertex must be 2π instead of the approximating each chart by a developable surface as in previous studies.

2. DEVELOPABLE SHAPES

When we express 3D figures by paper, the important thing is its developability. The developability means the property that can develop onto a plane without its own stretching. In this research, we call the figure that can be developed onto a plane without stretching *developable shape*. Developable surfaces, such as a cylinder, cone, and plane, are instances of developable shapes.

The general approach of making a figure's development is that the figure is segmented into some charts and then each chart is developed on the flat. In order to express a figure by paper, each chart is a single developable surface or consists of some various kinds of developable surfaces. Most of shapes can be separated into various kinds of surfaces, but it is quite rare that they consist of developable surfaces only.

Therefore it is necessary to approximate a target figure by developable shapes. On the other hand, we can argue that any mesh model consists of the developable surfaces because it is composed of large amounts of triangles. However it is not feasible that a mesh model is developed as a collection of original triangles, because it has a huge number of triangles.

Hoppe [1] and Garland et al [2] introduced methods of surface simplification. In these methods, the triangles that compose the mesh model are combined together while preserving shape features of the original model. These methods are able to reduce the number of triangles. However, the resulting shape looks no more smooth as the original and does not approximate it precisely. Some other methods approximating the mesh with developable surfaces like cones or cylinders, are proposed by Julius et al. [3], Shantz et al. [4], and Massarwi et al [5].

D-Charts introduced by Julius et al. [3] is the method that segments a mesh into quasi-developable charts. However these charts are not absolute developable surface, so it is impossible that the charts are developed on the flat without stretching. Shantz et al. [4] introduced a segmentation of a mesh model as well as an approximation by developable surfaces so that each chart can be unfolded onto the plane without stretching. However this method requires refinement of charts' boundaries because approximation by developable surfaces may result in no-intersection between neighboring charts. So after the mesh model approximation, chart boundaries are refined to intersect each other. This method may generate non-uniformly small charts, which make difficult gluing charts into a papercraft.

Mitani et al. [6] introduced a method that produces strip-based approximation segmentation of a mesh model. This method succeeds in generating a relatively smooth papercraft. However, strip-based segmentations are long and usually need more charts to achieve better approximation, so it is too difficult to cut and glue.

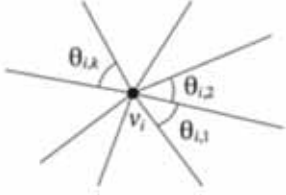


Figure 2: Developability around a Vertex.

3. PROPOSED METHOD

In general, a papercraft consists of some charts each of which can be unfolded onto a plane. The papercraft must accurately approximate the original mesh model while it must be easily assembled by gluing charts together. Over the last decade, most of studies have been aimed both at reducing the number of the charts and at simplifying of charts' shapes, because they claimed that the easiness of cutting and gluing depends on both number and shapes of charts. However, our method focuses only on reducing the number of charts because of two reasons. The first reason is that recent cutting plotters can automatically cut complicated charts. The second reason is that the easiness of gluing depends only on the number of charts. The correspondence of edges to be glued in a chart is trivial, because a pair of edges is glued together one by one along a slit.

Now, we consider origami as an example of the paper modeling. Origami can generate a complicated shape by creasing a sheet of paper. In other words, origami obtains flexibility of expression by introducing creases in a paper. We propose a segmentation method that introduces slits and creases into charts. By using creases, different developable surfaces can be combined into one chart. In addition, by using slits, it is expected that the approximation accuracy may be improved and gluing can be easy because the edge correspondence in a chart is trivial as we stated before. That is, due to creases and slits, the number of charts can be reduced and it can be easier to glue charts compared to previous methods.

However, it is difficult to introduce creases

and slits to a method based on developable surfaces like cones and cylinders as in previous research. Instead, we use the angle constraint around vertices [7]. This constraint states that a total angle around a developable vertex must be 2π (Figure 2),

$$2\pi - \sum_{k=1}^{m_i} \vartheta_{i,k} = 0, \quad (1)$$

where $\vartheta_{i,k}$ is the k -th incident sector angle of v_i . This angle constraint allows creases in charts. This constraint is used for chart segmentation as well as for developable deformation, which will be explained later in this paper. The developability based on the angle constraint can be realized by changing the coordinate of the vertices regardless of the mesh topology. Thus, the connectivity of charts can be achieved easily, which was the challenging issue in the method of [4].

3.1 FEATURE OF THIS METHOD

Our method achieves the following properties:

1. Each chart may contain creases within itself.
2. Each chart may have slits but their total length should be as short as possible.
3. The entire mesh model preserves its topology.
4. The user can control a number of charts.

The first property is mentioned previously. The second property is realized by leaving a less-developable vertex as a boundary vertex of a chart when a mesh model is segmented into charts. After chart segmentation, a certain part of a slit may be closed like a zipper if the closing is not critical for developability of the entire chart. This judgment is performed by taking account of the slits' length as well as on the angle constraint.

The third property is achieved by adjusting only positions of vertices in order that the entire model can be developable. Thus, boundary reconstruction, which is needed in [4], is no more necessary.

The last property is achieved by reducing the

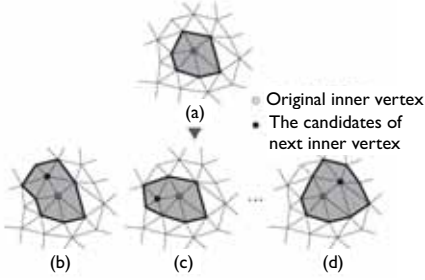


Figure 3: Vertices of Combined Charts.

number of charts. It is accomplished by repeatedly culling the smallest chart so that all charts are segmented to be equal-sized.

The properties 1 through 3 are new approaches that they are not attempted in previous research. They might achieve the significant reduction of number charts as well as improvement of approximation accuracy. The last property is also a new approach that user can control the number of charts.

3.2 ALGORITHM

There are two phases in our method, namely, mesh-segmentation phase (Section 3.2.1- 3.2.2) and deformation-and-development phase (Section 3.2.3-3.2.5). In the former phase, a mesh is segmented into charts each of which roughly satisfies the angle constraint. In the latter phase, all the charts are deformed so that they become developable by exactly satisfying the angle constraint. Figure.1 illustrates the entire algorithm of our method.

In the first step of initial segmentation, a mesh model is segmented into tentative charts, using the angle constraint around vertices (Figure 1(a)). This step will be explained in Section 3.2.1. Secondly, charts are selected by repeatedly culling the smallest chart so that every chart has almost the same number of triangles (Figure 1(b)). Section 3.2.2 will describe the detail of this step.

Thirdly, all the charts are deformed so that they become completely developable by moving vertices (Figure 1(c)). Section 3.2.3 shows

this step in detail. Fourthly, some parts of slits are closed like zippers if the closing is not critical for the developability of the entire charts (Figure 1(d)(e)). This step is expressed in Section 3.2.4. Finally, each chart is developed onto a plane and separated if a self-intersection occurs (Figure 1(f)). Section 3.2.5 illustrates the detail of this step.

3.2.1 Initial Segmentation

In this step, first we choose a primary vertex, called seed. Each chart grows from the seed that have the minimum cost of developability among vertices not belonging to any existing chart yet. The cost is defined by:

$$E(v_i) = \left| 2\pi - \sum_{k=1}^{m_i} \vartheta_{i,k} \right|, \quad (2)$$

which is based on the angle constraint specified in Equation (1). The initial chart consists of all triangles around the seed (Figure 3(a)).

There are two kinds of vertices that belong to a chart, namely, boundary vertex and inner vertex. A boundary vertex exists on a boundary of the chart. An inner vertex is not located on a boundary so that it is completely surrounded by triangles consisting of the chart. Therefore, the seed is the only inner vertex of the initial chart. A chart grows by changing one of boundary vertices into an inner vertex as shown in Figure 3. The boundary vertex is selected by evaluating the cost function defined as follows:

$$Cost(C) = \sum_{v_i \in V_I(C)} E(v_i). \quad (3)$$

Here, $V_I(C)$ is a set of the inner vertices of the chart C . Figure 3(b)-(d) are the candidates of the next grown charts from the initial chart shown in Figure 3(a). The next chart has the minimum cost among all the candidate charts. The chart grows until the cost of the chart $Cost(C)$ reaches a predefined threshold.

This step naturally produces slits in a chart. Figure 4 illustrates this process of emergence and growth of a slit. A slit occurs when one of boundary vertices is surrounded by triangles of the chart without being an inner vertex as shown in Figure 4(b). This type of boundary

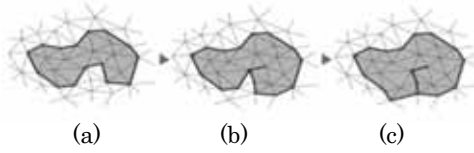


Figure 4: Emergence and Growth of a

vertex surrounded by chart triangles is excluded from the candidates of chart growing. Consequently, a slit is never eliminated and naturally grows like Figure 4(c).

Seed selection and chart growing are repeated until there exists no candidate of seeds. After this step, every vertices of the mesh model belong to one of charts. But some triangles may not belong to any chart especially around chart boundaries. Those triangles are incorporated into one of the neighboring charts so that the resulting boundary becomes the smoothest. Figure 1(a) shows an example of initial segmentation.

3.2.2 Chart Selection

The charts produced by the initial segmentation explained in Section 3.2.1 differ in size to a large degree. Some produced early in this step tend to be big, and some produced late in that tend to be small, because the early charts grows by taking highly developable vertices. The number of the chart is determined as a result according to the shape of the given mesh model and the threshold value of chart growing. This step reduces the charts to a user-specified number by selecting the seeds produced in the initial segmentation and re-segments the mesh model with the selected seeds.

Firstly, initial charts $C_i (i = 1, 2, \dots, n)$ are generated with the set of seeds obtained by the initial segmentation. Then One of the initial charts that has the lowest cost defined by Equation (3) is chosen, which means the chart has highest developability. The chosen chart is grown only once by changing one of its boundary vertices into its inner vertex. By repeating this process, namely growing the chart of the highest developability, growing

speed of a chart keeps the same pace with those of the other charts. This process is completed when all vertices on the mesh model belong to n charts.

The seed of the chart that have the least number of triangles is removed from the seed set. This means the number of seeds is decremented by one. Based on this new seed set, the mesh model is segmented again, which results in $n - 1$ charts. By repeating this process, the number of chart reaches to the user-specified number. Figure 1(b) shows a result of chart selection. Compared with the result of initial segmentation (Figure 1(a)), the chart number is significantly decreased and the sizes of chart become almost uniform.

3.2.3 Deformation For Developability

The charts obtained by the chart selection explained in Section 3.2.2 are nearly developable but not perfectly developable. This step calculates the coordinates of all vertices so that all charts become developable. That means all inner vertices satisfy Equation (1).

The coordinates of all vertices are denoted as below:

$$V = [x_1 y_1 z_1 x_2 y_2 z_2 \cdots x_n y_n z_n]^T,$$

where n is the number of the mesh model vertices. A set of functions corresponding to the inner vertices is defined as below:

$$F(V) = \begin{bmatrix} 2\pi - \sum_{k=1}^{m_1} \vartheta_{1,k} \\ 2\pi - \sum_{k=1}^{m_2} \vartheta_{2,k} \\ \vdots \\ 2\pi - \sum_{k=1}^{m_{n_i}} \vartheta_{n_i,k} \end{bmatrix}, \quad (4)$$

where n_i is the number of the inner vertices of all charts. Each function represents difference between 2π and the total angle around the inner vertex, that is it stands for the developability of the vertex. Thus, the charts become developable if V satisfies the following constraint:

$$F(V) = 0. \quad (5)$$

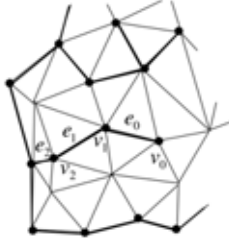


Figure 5 : Evaluation of Developability along a Slit.

All the charts of the mesh model turn to be developable by solving the linear equation system. It is achieved with the computation method proposed by Tachi[7]. It gives a feasible solution if the given charts are almost developable. This constraint is provided only for inner vertices and the boundary vertices are not constrained at all. Thus the boundary vertices can be freely moved in order that all the inner vertices satisfy the constraint. The neighboring charts sharing the boundary vertices preserve their connectivity though the vertices are moved. This means that the reevaluation of chart boundaries is not necessary. Figure 1(c) depicts a result of the deformation for developability. The curvature of the mesh model is concentrated at the slits or the boundaries.

3.2.4 Slit Closing

The chart at this point may have lots of slits that are produced during the chart selection process. These slits allow the charts to be highly developable. On the other hand, charts with less slits might be easier to be glued. This step closes certain parts of slits like zippers unless the closing is critical for the chart's developability. It is influenced not only by the angle around a vertex but also by the slits' length. The cost function representing the developability of a slit is defined as below:

$$E_s(v_i) = \sum_{j=1}^l (E(v_i) \sum_{j=1}^l \|e_j\|) \quad (6)$$

where $E(v_i)$ is developability cost of a vertex V_i defined by Equation (2) and $\|e_j\|$ is the



Figure 6 : A Self-Intersected Development.



Figure 7 : Separated Development .

length of an edge e_j . Thus, Equation (6) considers both the residual error of vertex developability and the length of the closing slit.

After the step of slit closing, the mesh model is no longer perfectly developable. Thus, the deformation explained in Section 3.2.3 is applied to the mesh model once again.

3.2.5 Chart Development

The last step of our method develops each chart onto a plane. Figure 6 illustrates an example of developed charts. Red and blue lines indicate valley folds and mounting folds, respectively. The folds are indicated if their folding angles are large enough. The threshold angle is $\pi/8$ in the case of Figure 6, which is a development of a part of the Stanford Bunny. It corresponds to the brown area of the Bunny's front leg shown in Figure 1(e). This step may result in self-intersection among developments as indicated with a violet circle in Figure 6. Those developments are separated into several pieces to resolve the self-intersection. Figure 7 is the

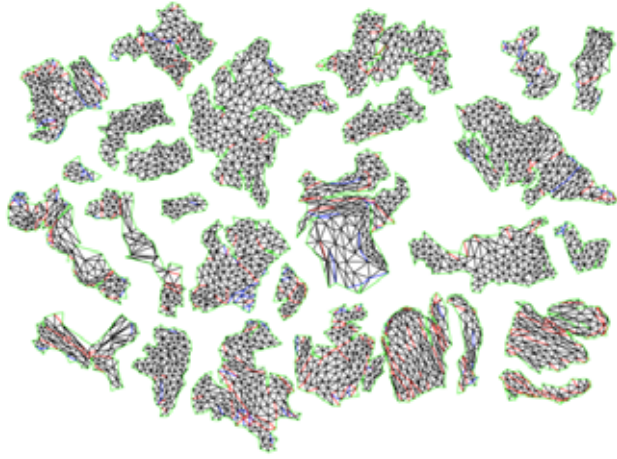


Figure 8 : An Example of Stanford Bunny's Development Making by the Proposal Method.



Figure 9 : An Example of Resulting Papercraft.



(a)



(b)



(c)

Figure 10 : Effects of the number of charts.

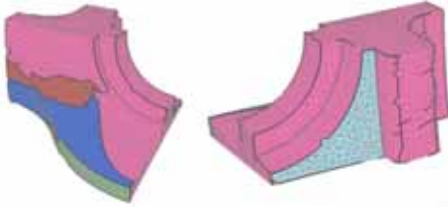


Figure 11 : An Example of Charts of the Fan Disk.

resulting developments by dividing the development shown in Figure 6. This separation was done manually.

4. RESULTS

Figure 8 shows sample developments created by the proposed method. They are developments of the Stanford Bunny consisting of 27 parts. It is fewer than the result of Shantz's method [4] that consists of 35 parts. Figure 9 depicts the resulting papercraft fabricated with the developments of Figure 8.

Figure 10 indicates the effect of chart selection. Figure 10(a) shows the result the charts in which are reduced to 5. Figure 10(b) and 10(c) shows those of 10 and 20, respectively. When reducing the number of charts, the accuracy of approximation is preserved by increasing the number of slits instead. Thus, the mesh model consisting of less number of charts may lose the approximation accuracy after the slit closing.

As the number of charts is fewer, the sizes of charts become bigger. This may result in self-intersection among charts. As a result, chart selection cannot control the final number of charts.

Figure 11 presents another sample of result generated by the proposed method. Machinery components like Fan Disk as shown in Figure 11 consist of planes and cylinders. It is expected to reduce the number of charts by connecting them with creases. The proposed method succeeded in reducing the number of charts of Fan Disk to only 5.

5. CONCLUSIONS AND FUTURE WORK

This paper introduces a method for developable mesh segmentation and approximation. Our method achieves the following properties:

1. Each chart may contain creases within itself.
2. Each chart may have slits but their total length should be as short as possible.
3. The entire mesh model preserves its topology.
4. The user can control a number of charts.

The properties 1 and 2 are advantageous for making papercraft because of the following reasons. Firstly, the original model can be well approximated. Secondly, the number of charts becomes small. The charts with creases and slits achieve better approximation with less number of charts than those of developable surfaces with simple boundaries. The property 3 frees from calculating chart boundaries. The property 4 allows the user to control the easiness of assembly, though there exists trade-off between approximation accuracy and the number of charts. Besides, when the charts are small in number, each chart size will become big, so finally the more self-intersection in each charts are came into being.

Fig. 11 presents another segmentation results of this algorithm. Machinery component like Fan Disk like Fig.11 is often consists of some planes and cylinders, then it is expected to reduce the number of charts by using our method that is with creases and slits. In the case of Fig.11, Fan Disk is composed of only 5 charts.

In the future, we want to improve segmentation step in this method. After development, some charts may intersect itself. By improving the segmentation step, we can resolve this self-intersection.

ACKNOWLEDGMENTS

This work was partially supported by JSPS Grant-in-Aid for Scientific Research (B) (22300030).

REFERENCES

- [1] Hoppe, H., “Progressive Meshes”, SIGGRAPH 96 Conference Proceedings, ACM SIGGRAPH (1996), pp. 99–108.
- [2] Garland, M. and Heckbert, P. S., “Surface simplification using quadric error metrics”, SIGGRAPH 97 Conference Proceedings, ACM SIGGRAPH (1997), pp. 209–216.
- [3] Julius, D., Kraevoy, V. and Sheffer, A., “D-Charts: Quasi-Developable Mesh Segmentation”, Computer Graphics Forum, Vol. 24, No. 3 (2005), pp. 581–590.
- [4] Shatz, I., Tal, A. and Leifman, G., “Papercraft models from meshes”, Visual Computer, Vol. 22, No. 9 (2006), pp. 825–834.
- [5] Massarwi, F., Gotsman, C. and Elber, G., “Papercraft Models using Generalized Cylinders”, Proceedings of the 15th Pacific Conference on Computer Graphics and Applications (2007), pp.148–157.
- [6] Mitani, J. and Suzuki, H., “Making papercraft toys from meshes using strip-based approximate unfolding”, ACM Transactions on Graphics, Vol.23, No. 3 (2004), pp. 259–263.
- [7] Tachi, T., “Freeform Variations of Origami”, Journal for Geometry and Graphics, Vol. 14, No.2 (2010), pp. 203–215.

ABOUT THE AUTHORS

1. Kuniko Takahashi is a systems engineer at Fuji Xerox. She received her Master’s degree from the University of Tokyo in 2013. Her research interests lie in computer graphics including papercraft. Her e-mail address is kunikot@me.com.
2. Tomohiro Tachi is an assistant professor of general systems studies in the University of Tokyo. His research interests include origami, computational design, and digital fabrication.
3. Yasushi Yamaguchi, Dr.(Eng.), is a professor at the University of Tokyo. His research interests lie in image processing, computer graphics and visual illusion, including visual cryptography, computer aided geometric design, volume visualization and painterly rendering. He has been serving as a president of JS GS and a vice president of ISGG. His email His e-mail address is yama@graco.c.u-tokyo.ac.jp.

Arm-Z: a modular virtual manipulator

Machi ZAWIDZKI and Takehiko NAGAKURA
Massachusetts Institute of Technology, USA

ABSTRACT: This paper presents Arm-Z - a concept of a kinematic system composed of congruent modules and capable of complex movements. Three fundamental spatial movements of the modular arm are explained and illustrated: *extension*, *translation*, and *flexure*.

Keywords: Flexible arm, Free-form, Pipe-Z, Arm-Z, Modular robot, Mathematical knot

1. INTRODUCTION

Pipe-Z is a parametric design system introduced in [1] which comprised of congruent modules (PZM) allows the creation of complex three-dimensional single-branch structures, for example mathematical knots. Fig. 1 shows Trefoil and Cinquefoil realized by PZ_{12} , that is a Pipe-Z based on dodecagonal PZMs. This paper presents the concept of modular Arm-Z (AZ), which is an extension of PZ. AZ which is also constructed from a sequence of PZMs introduces dynamic transformations to the previous system. AZ is constructed by assembling PZMs in a sequence, so the top face of the previous unit becomes the base for the next one. The successive unit i is rotated by twist angle κ_i , which can have real or discrete values. In the latter case such rotations are denoted by k_i . In the further examples PZMs are based on hexagon ($n = 6$) and the entire configuration is denoted as AZ_6 . In AZ_6 a subsequent unit can be added at six rational (dihedral) angles, so the facets of adjacent units are aligned. In further examples the following PZM parameters have been selected: $r = 0.25$, $\zeta = 22.5^\circ$, $s = 0.4$, and $n = 6$, as illustrated in Fig. 2. AZ is capable of performing complex spatial movements by controlling solely the sequence of twist angles k_i , in other words by the relative twists of PZMs. The twist angles which determine the spatial configuration of AZ are discrete. Therefore the changes of entire AZ are also discrete and occur at discrete time-steps. As a consequence,

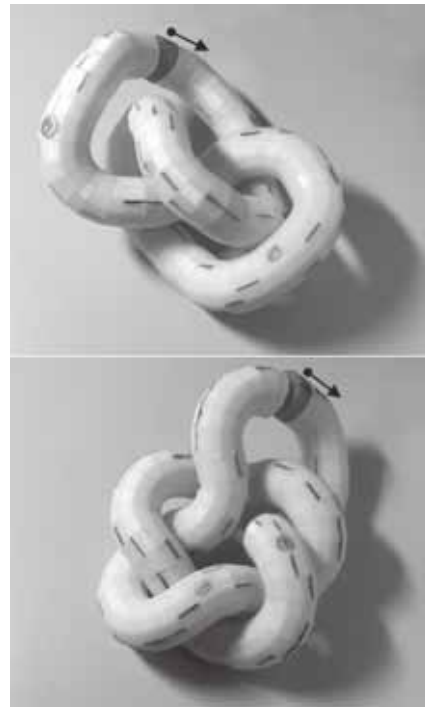


Figure 1: Two mathematical knots realized by PZ_{12} . Trefoil and Cinquefoil are shown on the top and bottom, respectively. The initial PZM and the PZ_{12} direction are indicated in gray and by gray arrow, respectively.

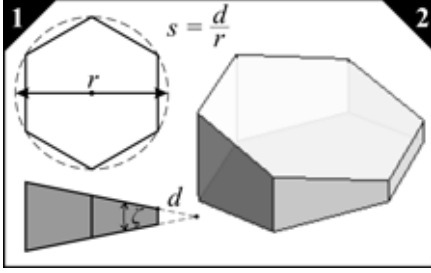


Figure 2: PZM_6 : 1) On the top and bottom: the top and side views, respectively. 2) An axonometric view.

at one time-step any number of twist angles k_i can change by at most one dihedral angle. In the following subsections three fundamental spatial operations are explained and illustrated: *extension*, *translation*, and *flexure*. For an interactive demonstration illustrating these movements see [2]. In all the examples the first PZM is set arbitrarily and shown as semi-transparent. The coloration of PZMs follows the right-hand rule, that is: the PZMs which will turn right (along the arm) and left are shown in green and red, respectively.

2. EXTENSION

Extension is the simplest movement, defined here as a “continuous” elongation of a part of AZ (ideally without any other displacements). Fig. 3 shows extensions at two consecutive time-steps. The history of changes of the AZ shape can be represented by a matrix (MOC). Following the common convention in discrete heuristic optimization, the sequence of units can be considered a *genotype*; positions in the AZ sequence are called *loci* and the possible alterations at the given *locus*: *allele*. Here *allele*: $(-1, 0, 1)$ stand for a dihedral twist to the left, no change, and dihedral twist to the right, respectively. The colors at each *locus* correspond the coloration introduced previously, therefore $(-1, 0, 1)$ are represented by red, white and green, respectively. Figure 4 shows the MOC corresponding to Fig 3.

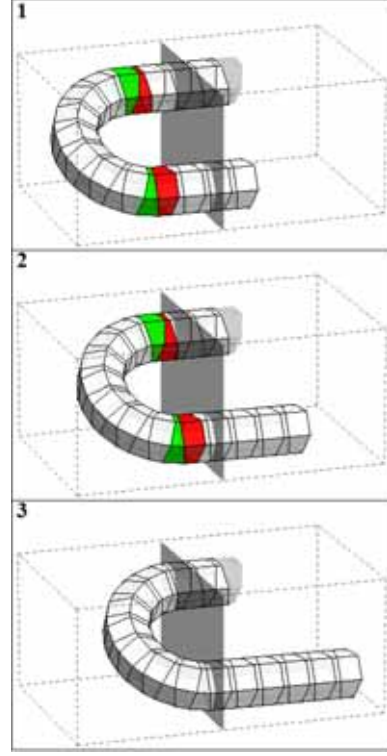


Figure 3: Extension: The gray plane is introduced for clarity. The segment on the right side becomes longer; the changes to the left part of AZ are minor - the general shape is maintained.

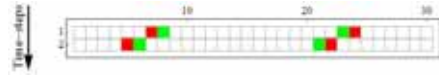


Figure 4: MOC corresponding to Fig. 3. Each time-step is represented by a row.

3. TRANSLATION

Translation is a “continuous” displacement of a distinct PZM configuration (which ideally does not change) in a given direction along AZ. Here a vertical “pipe” is displaced along a horizontal “pipe”. This transformation is more problematic

than *extension* as shown in Fig. 5. Figure 6 shows the MOC corresponding to Fig. 5.

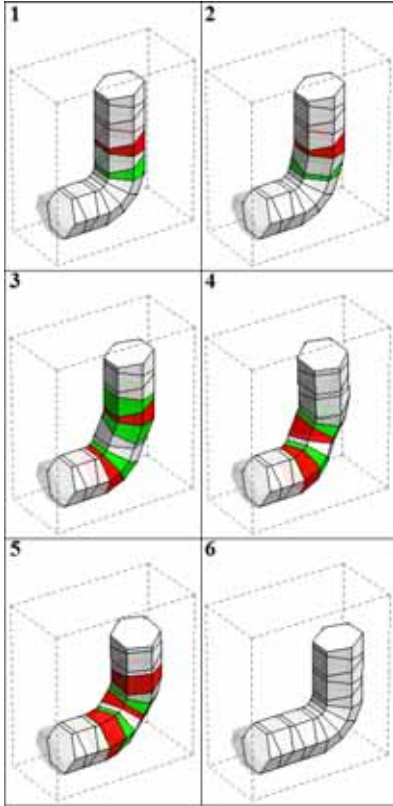


Figure 5: The vertical “pipe” (VP) moves along the horizontal “pipe”. 2-3). VP becomes eccentric; 4-5) VP is eccentric and slightly nonvertical.

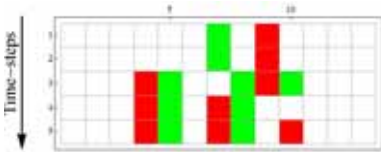


Figure 6: MOC corresponding to Fig. 5.

4. FLEXURE

Flexure is defined here as a “continuous” change of shape from a straight segment of given number of PZMs to a full circle of the same number of PZMs. It is the most difficult transition as shown in Fig. 8. Figure 7 shows the MOC corresponding to Fig. 8.

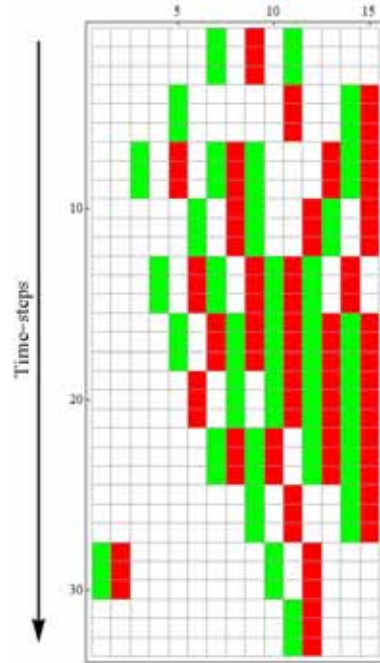


Figure 7: MOC corresponding to Fig. 8.

The transition shown in Fig. 8 has been composed semi-manually. Although most likely it is not optimal, it illustrates that AZ has potential of performing complex spatial movements rather efficiently.

5. CONCLUSIONS

Arm-Z (AZ) reminds of some existing modular systems based on congruent units, e.g.: so-called “tangle toys” and modular snake robots (Mod-snakes) [4]. There are however major differences.

Unlike the first system which realizes various embeddings of the same *Unknot*, AZ is capable of constructing any knot. The challenges for Mod-snakes are also different - they focus on robustness and mobility [7]. AZ also reminds of the robotic endoscopes [5], which however are not modular (in the same sense); the geometric tasks described here have the same nature as the minimally invasive surgery [6]. This paper, however, demonstrates that a modular system whose units have only one degree of freedom is capable of practically any 3D motion. *Extension*, *translation*, and *flexure* have been executed rather satisfactorily. AZ is therefore capable of performing complex spatial tasks by combination of these fundamental movements. The first movement is relatively simple, however, controlling the extension speed needs further study. The second and third movements are difficult and very difficult, respectively. Although the example of *translation* might not be optimal - it is acceptable. However, the example of *flexure* is quite problematic - the eccentricity and other kinds of deformations are evident at many steps of the transition. Nevertheless, an AZ system based on units with larger number of sides would be capable of smoother transition in all kinds of movements.

The illustrative examples presented here have been composed semi-manually. Development of more practical control method is necessary. Evolutionary algorithm (EA) is a possible approach as suggested in [3].

ACKNOWLEDGMENTS

This work was completed as part of the Singapore University of Technology & Design and Massachusetts Institute of Technology Postdoctoral Fellows Program (SUTD-MIT PDP).

REFERENCES

REFERENCES

[1] M. Zawidzki and K. Nishinari. Modular Pipe-Z System for Three-Dimensional Knots. *Journal for Geometry and Graphics*, 17(1): 81–87, 2013.

[2] M. Zawidzki. Arm Manipulations. Wolfram Demonstrations Project, URL: <http://demonstrations.wolfram.com/ArmManipulations/>, 2014.

[3] M. Zawidzki and K. Nishinari. Preliminary Results for Pipe-Z Knot Optimization. in Y. Tsompanakis, (Editor), *Proceedings of the Third International Conference on Soft Computing Technology in Civil, Structural and Environmental Engineering*, Civil-Comp Press, Stirlingshire, UK, Paper 29, 2013.

[4] C. Wright A. Johnson, A. Peck, Z. McCord, A. Naaktgeboren, P. Gianfortoni, M. Gonzalez-Rivero, R. Hatton, and H. Choset. Design of a modular snake robot. In *Intelligent Robots and Systems, 2007. IROS 2007. IEEE/RSJ International Conference on* ., 2609–2614., 2007.

[5] Y. Chen, J.M. Oliveira and I.W. Hunter. Two-axis bend sensor design, kinematics and control for a continuum robotic endoscope. *Robotics and Automation (ICRA), 2013 IEEE International Conference on* ., 704–710, 2013.

[6] R.J. Murphy, M.S. Moses, M.D.M. Kutzer, G.S. Chirikjian and M. Armand. Constrained workspace generation for snake-like manipulators with applications to minimally invasive surgery. *Robotics and Automation (ICRA), 2013 IEEE International Conference on* ., 5341–5347, 2013.

[7] M. Yim, K. Roufas, D. Duff, Y. Zhang, C. Eldershaw and S. Homans. Modular reconfigurable robots in space applications. *Autonomous Robots*, 14(2-3): 225-237, 2003.

ABOUT THE AUTHORS

1. Machi Zawidzki, D.Eng., M.Arch. Dr. Machi Zawidzki is currently a postdoctoral fellow at the Department of Architecture of Massachusetts Institute of Technology, funded by the Singapore University of Technology and Design Postdoctoral Program. The research is titled: "Effective computational methods for grid and raster-based modeling of practical problems in architectural and urban layout design". In 2013 he has completed a postdoctoral research at the Research Center for Advanced Science and Technology of the University of Tokyo, funded by a grant from the Japan Society for the Promotion of Science. The research was titled: "Improvements of the Seniors' Quality of Life through Application of Innovative Computational Systems to the problems of Accessibility, Ergonomics and Housing & Living Environment". He received Eng. and M.Arch degrees from Warsaw University of Technology, Poland, in 2000 and the D.Eng degree from Ritsumeikan University, Japan, in 2010. From 2002 to 2005 he worked at Howard Alan Architects, Chicago, a studio specializing in sustainable and in particular - solar architecture. From 2006 to 2007, he worked at Studio E Architects, London, England, a studio specializing in sustainable and public architecture.
2. Takehiko Nagakura, PhD. Professor Nagakura is an architect from Tokyo. At MIT, he teaches courses related computer-aided design, and his research focuses on the representation and computation of architectural space and formal design knowledge. He has founded and led Architecture, Representation and Computation group (ARC) since 1996. His completed building, Gushikawa Orchid Center in Okinawa, Japan has been awarded with SD Review Award (1998) and Nikkei Kyushyu District New Office Award (1999). He is the co-author of *Gendai*

Kenchiku no Hassou (Ideas in Contemporary Architecture)(Maruzen, 1989) and translated William J. Mitchell's *the Logic of Architecture* (MIT Press, 1990) into Japanese (Kajima, 1991). His essays include "Shape Recognition and Transformation" in the *Electronic Design Studio*, edited by William J. Mitchell, Patrick Purcell and Malcolm McCullough (MIT Press, 1990). The result of his past MIT studio courses in electronic design environment is summarized in a co-authored article, "Digital Pinup Board – The Story of the Virtual Village Project" in *Virtual Design Studio*, edited by Jerzy Wojtowicz (Hong Kong Press, 1995). Before coming to MIT in 1993, he worked for Fumihiko Maki in Tokyo, and was an instructor at Harvard University, Graduate School of Design. He earned Bachelor of Engineering in Architecture from Tokyo University in 1985, Master of Architecture from Harvard University in 1987, Master of Engineering in Architecture from Tokyo University in 1988, and completed his PhD at Harvard in 1996. In 1985 he received the prestigious Ishizaka Memorial Foundation scholarship from the Japanese Federation of Economic Institutions. He is the recipient of the Japan Information Culture Society Grand Prize in 1999.



Figure 8: A possible transition from straight segment to complete circle.

ASSESSMENT METHOD FOR AN EDGE ALIGNMENT FREE HYBRID IMAGE

Sripian PEERAYA¹ and Yamaguchi YASUSHI²

¹King Mongkut's University of Technology Thonburi / JST CREST, Thailand

²The University of Tokyo / JST CREST, Japan

ABSTRACT: A hybrid image allows multiple image interpretations to be modulated by the viewing distance. It can be constructed based on multiscale perceptual mechanisms of the human visual system, by combining low and high spatial frequencies of two different images. The original hybrid image synthesis method limited to similar shapes of source images that were aligned in edges only. Our research in [10] proposed two methods for synthesizing a hybrid image from dissimilar shape images or unaligned images with the use of noises and complementary color, so called "Edge-alignment free hybrid image". However, there was no measurement method for actual perception of the low or high spatial frequency in a hybrid image. To ascertain that these two approaches successfully separate the spatial frequency at each viewing distance, a hybrid image's main property must be proved valid. That is, it must present only the high frequency image at a close viewing distance and only the low frequency image from far away. To measure this property, this paper introduces four significant probabilities of different hybrid image recognition type. With the experiment based on the proposed assessment method, we found that our two synthesis methods scored best in terms of keeping a hybrid image's main property.

Keywords: Contrast sensitivity, Human perception, Hybrid image, Scale space, Vision experiment

1. INTRODUCTION

Hybrid image is considered as a kind of optical illusion called "double image", in which two kinds of information can be derived from a single image representation. This kind of image is often used in the field of psychology as a tool to experiment. Oliva et al. introduced hybrid image in [6] based on series of experiments that used hybrid stimuli as an experimental tool to study human visual system.

Hybrid image exploits multi-scale perceptual mechanisms of human visual system to present a particular spatial frequency information in the image at a certain viewing scale, allowing multiple image interpretations to be modulated by the viewing distances. In order to see another information that is hidden in the hybrid image, the changing of degree of visual angle is required.

The degree of visual angle is relative to actual viewing distance. Hybrid image, I_H , is obtained by combining two images, I_1 and I_2 , at different spatial scales as below:

$$I_H = L_p(I_1) + H_p(I_2),$$

where L_p is a low-pass filter and H_p is a high-pass filter.

2. BACKGROUND

The sequence of spatial frequency in visual recognition is one of many unsunsumarized question. Traditional experimental procedures include the use of masking in an image, such as "critical band masking", that deteriorates the information in an image by selectively affecting the critical band. It includes filtering out a particular spatial frequency band from the image, or overlaying spatial frequency noises on the image.

In 1994, Schyns and Oliva introduced a new experimental paradigm that allowed the perception of two different interpretations based on spatial frequency bands in one image, a hybrid stimulus. A hybrid stimulus is composed of the low frequency band and the high frequency band of two different images. It is a kind of critical band masking that use meaningful image as a mask instead of noises. With the use of hybrid stimuli in the experiment, it is possible to measure which frequency band was attended at which stage of recognition [7]. Subjects were shown the hybrid stimuli for short exposure condition (30 ms) and long exposure condition (150 ms). The participants' task was to identify whether or not the shown stimulus contained a subsequent normal image that corresponded to the high or low frequency image. Experimental results were analyzed by measuring the proportion of correct hit for each frequency image. As expected, hybrids were preferentially categorized in a coarse-before-fine sequence, that is, the brief presentation of the stimulus resulted in the observer matching coarse structure more correctly. Meanwhile, longer presentation of the same stimulus resulted in bias toward fine structures. This research showed that the proportion of correct in the perception of the high spatial frequency, or the low spatial frequency of an image was used as recognition measurement. The correct or incorrect hit for one spatial frequency did not indicate whether another spatial frequency image was perceived simultaneously or not.

The later experiment, which studied scale perception in flexible scene categorization, considered the simultaneous perception of both spatial frequencies in a hybrid stimulus [4, 8]. The work in [8] tested whether categorization task effects the preferential selectivity of spatial frequency band by the use of face stimuli. Again, they confirmed in the debriefing phase that the observers perceived only one spatial frequency from a single hybrid image. Several observers mentioned that they perceived the hybrid of two faces. Similar to the experiment with diagnostic scene, the

recognition of two faces in one stimulus was considered as the failure to separate the perception of spatial frequency in the experiment. Since the main goal of the experiment was to measure which spatial frequency was used during the task, the perception of both spatial frequencies could induce a bias. The subjects were forced to select the information from one spatial frequency in the case that both spatial frequencies were present. Therefore, they could strategically choose to only report the face associated with one of the two scales. This study discarded an entire subject's data if he/she observed both spatial frequencies simultaneously during the experiment.

2.1 Hybrid image

In order to create a compelling hybrid image with the original hybrid image synthesis method proposed in [6], the source of both images should follow the rule of perceptual grouping, that is, the source images should be aligned in edges to prevent overlapping. To achieve this, some research effort emphasized on the search of an optimal pairing image that will align well with the main image, as in [11]. When composing two images that are not aligned in edges, parts of the low frequency image that are not masked by the high frequency image form a strong perceptual grouping cue even when seen up close. Consequently, not only information from the high spatial frequency, but also from the low spatial frequency could be conveyed in a single glance when looking at the hybrid image closely.

Our research intended to create a hybrid image that does not require the alignment of edges in both image sources, namely, *edge-alignment free hybrid image*. Ideally, the hybrid image that the low frequency image is perceived as noise, or disregarded completely when viewing closely. In the creation of an edge-alignment free hybrid image, we focused on the presentation of the high frequency image, when viewed from a short distance. In [10], we attempted to make the low frequency image less noticeable, while make the high frequency image become more noticeable

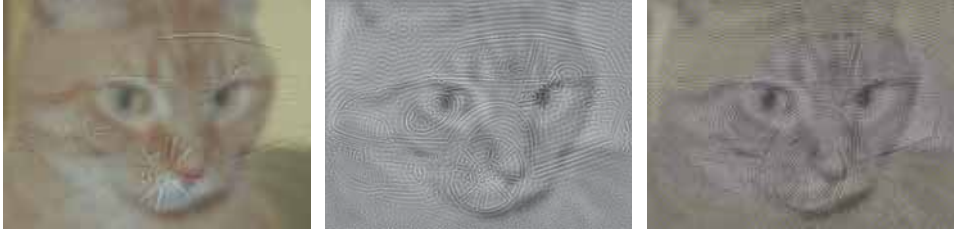


Figure 1: Hybrid images by a) [6], b) Noise-inserted method [10] and c) Color-inserted method [10]

at the same time by introducing *noise-inserted method* and *color-inserted method* accordingly. The idea of using noise in a hybrid image is originally from the previous work by [3]. They applied additional noises into the high frequency image to hide the non-overlapping parts of the low frequency image when viewing from a close distance. In [9, 10], we proposed some methods that prevent the resulting hybrid image from being too ambiguous caused by too much noise, which is accomplished by adding and enhancing noises while referring to both source images' edge location. Meanwhile, the color-inserted method uses colors to emphasize the high frequency image when viewed up close, while not interrupting the perception of the low frequency image when viewed from far away. To achieve this, we use chromatic sinusoidal gratings. The example of hybrid images are shown in Figure 1

3. PROPOSED ASSESSMENT METHOD

Our research is motivated by the lack of measurement procedure for spatial frequencies perception in a hybrid image. This work aims at providing the assessment methods for edge-alignment free hybrid images. To measure the success of a method for creating visual illusion, it is common to look at the actual effect of the result image. Accordingly, we measure the quality of hybrid image synthesizing methods by evaluating the effect produced by the final result, the interchanging of the perceived image by viewing distances. A good hybrid image must present only the high frequency image when viewed closely, and only the

low frequency image when seen from far away. This property is called *the separation of spatial frequency*. We consider a hybrid image ineffective when both spatial frequencies are perceived simultaneously at such viewing distances. The following sections describe a suitable experimental method and analysis procedure that tests the efficiency of a hybrid image with regard to the separation of spatial frequency.

3.1 Separation of spatial frequency

To find which hybrid image performs better in separation of spatial frequencies at a particular distance, we use four types of seeing probability:

- $P[L, H]$: both spatial frequencies;
- $P[L, \bar{H}]$: only the low frequency image;
- $P[\bar{L}, H]$: only the high frequency image; and
- $P[\bar{L}, \bar{H}]$: none of the spatial frequencies;

with the constraint as follows:

$$P[L, H] + P[L, \bar{H}] + P[\bar{L}, H] + P[\bar{L}, \bar{H}] = 1. \quad (1)$$

Although the measurement should be made with hybrids of natural images, it is costly to create much variations of stimuli and there existed some biases for interpreting image categories. To avoid such difficulties, a single-digit image is used for each spatial frequency. Examples of stimuli are shown in Figure 2. A hybrid stimulus is composed of two different numbers, each for the high or low frequency image. This paper proposes and

tests this evaluation method using non forced-choice experimental procedure with reasons described later.

3.2 Experimental procedure

In order to measure whether both spatial frequencies were perceived simultaneously, the observer must provide answers to both spatial frequencies. We considered the forced-choice and non forced-choice experimental procedures for this task. The forced-choice procedure forces the observer to provide answers for both spatial frequencies even if only one frequency is perceived, while the non forced-choice procedure allows observers to respond to only what they really recognize from a given image, including total skipping. We conducted a small experiment with four students to decide which procedure provided the most promising result. Observers commented that the forced-choice procedure was less practical than non forced-choice procedure because they felt that it was unnatural to guess all the answers of what they did not really see. Also, the experiment would require a high degree of concentration for a specific amount of time due to a large set of stimuli; therefore, the observers favored being able to skip unclear images. Although providing a skip option is not common in psychophysical procedures, it would be better to adopt an experimental procedure that measures actual recognition. In addition, forcing the observers to provide answers to both spatial frequencies even though only one spatial frequency is observed, can result in bias. In [5, 8], the experimenters had to discard the data from the observers who mentioned perceiving both spatial frequencies during the experiment, because the observers may report only a specific spatial frequency or strategically decide to report only a specific number for an uncertain answer throughout the entire experiment. Therefore, it is better to provide observers the option of skipping questions. We selected the non forced-choice procedure as an assessment procedure for our proposed hybrid synthesis method.

For a given image, observers may give answers

in three different ways, answers for both spatial frequencies, answer for only one spatial frequency, or not to provide any answer at all. Observer’s given answers are matched with the right answer to find whether the corresponding spatial frequency, low “L” or high “H” is correctly reported. If the answer does not match with any spatial frequency, it will be denoted with “E”. The observer may choose to give answer for only one spatial frequency, in this case, “S” stands for skipping an answer for the particular spatial frequency. Table 1 shows 8 possible answering patterns and their corresponding probabilities. ✓ and ✗ is used for cases that any spatial frequency is answered correctly and incorrectly.

Table 1: Observer’s answering pattern

	Low freq.	High freq.
$P[\text{LH}]$	✓	✓
$P[\text{LE}]$	✓	✗
$P[\text{EH}]$	✗	✓
$P[\text{EE}]$	✗	✗
$P[\text{LS}]$	✓	Skip
$P[\text{SH}]$	Skip	✓
$P[\text{ES}]$	✗	Skip
	Skip	✗
$P[\text{SS}]$	Skip	Skip

At each viewing distance, these probabilities for answered patterns add up to one for each hybrid image synthesis method, thus the degree of freedom of the probabilities for answered pattern is seven. We assume that some observers try to answer the spatial frequencies even if the observers did not see them. In other words, some people tend to guess situations where they are uncertain. Conversely, some might not answer the question unless they are sure of what they see, instead choosing to skip questions when they are unsure of the answers. To avoid personality bias, we introduce a factor called trial rate $t[X, x]$. Here, X denotes the type of spatial frequency that is not actually seen as “L,” “H” or “LH,” and x denotes how many images an observer attempts to guess.

For instance, $t[H, 1]$ indicates the rate at which the observer tried to guess an answer for the high frequency image, even if it is not properly seen, while $t[H, 0]$ depicts the rate that an observer did not attempt guessing. Likewise, $t[L, x]$ indicates whether an observer tried to guess an answer for the low frequency image ($x = 1$) or not ($x = 0$). In cases where both spatial frequencies are unclear to the observer, $t[LH, x]$ represents the rate that an observer attempts to answer x spatial frequencies. There are three constraints among the trial rates as below:

$$\begin{aligned} t[H, 1] + t[H, 0] &= 1, \\ t[L, 1] + t[L, 0] &= 1, \\ t[LH, 2] + t[LH, 1] + t[LH, 0] &= 1. \end{aligned} \quad (2)$$

Therefore, the degree of freedom of trial rates is 4 ($=7-3$).

The probability of guessing correctly must be considered when the observer guesses uncertain digits. We define it as guess rate $g[X, Y]$, where X stands for types of spatial frequency attempted guessing (L, H, LH), and Y stands for the result of the guess. Correctly guessed spatial frequencies are labeled “L” and “H” for low and high frequencies, respectively. “E” is used to indicate incorrect guesses. When the observer answers both special frequencies, guess rates, $g[X, Y]$, can be pre-computed as follows:

$$\begin{aligned} g[L, L] &= g[H, H] = \frac{1C_1}{9C_1} = \frac{1}{9}, \\ g[L, E] &= g[H, E] = \frac{8C_1}{9C_1} = \frac{8}{9}, \\ g[LH, LH] &= \frac{2C_2}{10C_2} = \frac{1}{45}, \\ g[LH, LE] &= g[LH, EH] = \frac{1C_1 \ 8C_1}{10C_2} = \frac{8}{45}, \\ g[LH, EE] &= \frac{8C_2}{10C_2} = \frac{28}{45}. \end{aligned} \quad (3)$$

For example, $g[L, L]$ denotes the rate of correct low frequency guesses out of the remaining nine digits. Thus, its probability is one in nine. Also, $g[LH, LE]$ is the guess rate of correctly guessing only the low frequency from both spatial frequencies. Since there are 45 possible combinations

of digits, the probability is 28 in 45. When an observer attempts only one out of two uncertain frequencies, guess rates are computed out of ten possible digits as below:

$$\begin{aligned} g[LH, L] &= g[LH, H] = \frac{1C_1}{10C_1} = \frac{1}{10}, \\ g[LH, E] &= \frac{8C_1}{10C_1} = \frac{8}{10}. \end{aligned} \quad (4)$$

For instance, $g[LH, L]$ and $g[LH, H]$ denote that an observer skipped one frequency and guessed the other frequency correctly, low or high accordingly. The probability is one out of all possible ten digits. Meanwhile, the probability of guessing the frequency incorrectly, $g[LH, E]$, is eight in ten. The calculation of probabilities are explained in Appendix.

4. EXPERIMENT

This experiment’s objective was to show that our proposed hybrid synthesizing methods successfully separated the spatial frequency over other previous methods. The probabilities of seeing spatial frequencies introduced in Section 3.1 are used as measures in this experiment. Several pilot studies were done to determine suitable stimulus exposure times. Viewing distances were randomized to avoid the effect of image presentation order as indicated in [1]. To achieve that, image sizes were used to simulate viewing distances instead of physical distances.

4.1 Subjects

Fourteen observers (7 males, 7 females, age range 19–30), participated in our experiment. Each reported normal or corrected-to-normal vision. All were tested for color vision deficiencies with Ishihara color test patches, and none were found to be colorblind. All subjects gave informed consent.

4.2 Stimuli

Four types of stimuli were constructed from four different hybrid image synthesis methods respectively. Viewing distances were simulated by image sizes. There were seven viewing distances, almost equally spaced by degree of visual angle, which are 2.43° , 3.81° , 5.12° , 6.54° , 7.8° , 9.26°

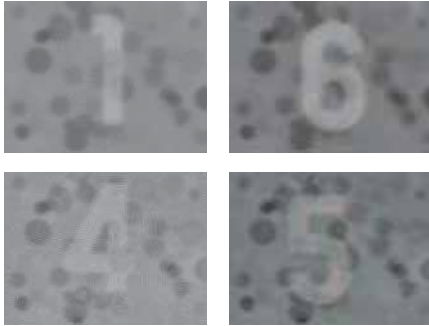


Figure 2: Samples of stimuli for our experiment. A hybrid image composed by the original method [6] is shown in top left, Konishi’s method [3] in top right, and our noise-inserted and color-inserted methods at the bottom left and bottom right, respectively.

and 10.58° , equivalent to 305 cm, 195 cm, 145 cm, 115 cm, 95 cm, 80 cm and 70 cm, relatively.

Hybrid images are composed by the original method [6], the method proposed by Konishi et al. [3], and our noise-inserted and color-inserted methods as shown in Figure 2. In this experiment, we referred to these methods as **Oliva**, **Konishi**, **Noise** and **Color**. Fifty images are constructed with a hybrid synthesis method for a viewing distance, represented by image’s size. Source images were resized prior to the synthesis of the hybrid image, and the cutoff frequency was calculated according to the final image size.

The high and low frequency image was an image of a single-digit number. A hybrid image was generated with the restriction that the numbers in both spatial frequencies must be different to prevent ambiguity in perception. Although the number was randomly positioned in each image, both images were created so that the numbers overlapped each other in some way. All images were displayed in the center of the screen without any apparent visual disturbances in the viewing angle.

4.3 Procedure

The experiment was composed of 1400 hybrid images in total. We divided the entire experiment into five blocks, each block composed of 280 hybrid images. Hybrid image types and sizes were randomly ordered in the presentation. Each image was presented on ColorEdge CG243W (IPS LCD color-managed 24.1-inch screen monitor), with a resolution of 1920×1200 pixels. The color profile of the monitor used in calculation of the grating gamut was created by a Konica Minolta CS-200 Chroma meter.

A hybrid image was displayed for 190 ms, followed by two consecutive answer boxes on the screen. Before proceeding to the next hybrid image, a noisy image was displayed for 500 ms as an inter-stimulus interval to prevent an afterimage. Figure 3 illustrates the experimental procedure. The observer was instructed to look at the image and to type the numbers in the answer boxes. Each answer box was designed to hold a single character. The observer could enter a number or a symbol that allowed him/her to skip the question. If he/she input the symbol for skipping on the first answer box, it would mean that he/she skipped the entire image. Thus, two answer boxes were displayed consecutively only in the case that at least one spatial frequency was answered. The order of answers did not affect the analysis process when the observer answered two numbers. There was no time constraint during answering phase. We used a chin rest to maintain the viewing distance between the observer and the monitor. The experiment was done under a dim lighting environment. The experimenter stayed with subjects to record occurrences of typing errors. There was a practice session for the observers to practice the use of the numerical pad on the keyboard before entering the experiment. Data from this session was not used in the results analysis.

5. RESULTS AND DISCUSSION

Four probabilities of seeing spatial frequencies were calculated with the method proposed in Sec-

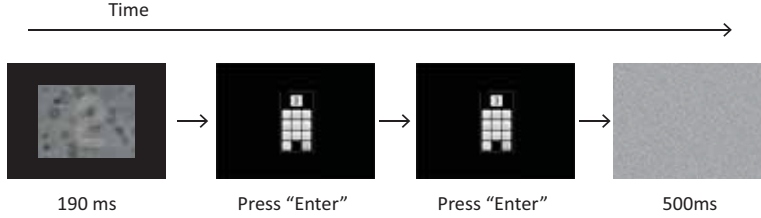


Figure 3: The procedure of our experiment.

tion 3.2. We compared the probabilities among all hybrid image synthesis methods. In this section, we describe findings from the experiment with some supporting statistical evidence.

5.1 Image separation performance

Our aim is to prove that the proposed methods perform better in spatial frequency separation. At all viewing distances, a good hybrid image should present only one spatial frequency. Therefore, to compare image separation performance, the integral of $P[L, H]$ of all viewing distances for each image type is used.

$P[L, H]$ is derived by Equations (1) and (7). Since the probability of seeing either a high or low frequency image is related to the viewing distance, we can model it as a psychometric function of distance x . To integrate $P[L, H]$ values of all viewing distances, $P[L, H]$ values are fitted by a function with regard to a viewing distance with a psychometric function in the MATLAB Palamedes toolbox [2]. We found that the cumulative normal distribution function (Equation (8) in Appendix) of viewing angle instead of distance was the best to fit the data, despite its negative x domain.

Each observer's $P[\bar{L}, \bar{H}]$ for any type of method is negligibly small at all distances. $P[\bar{L}, \bar{H}]t[LH, x]$ was also negligibly small, which satisfies the constraints for solving Equation (7). Accordingly, $P[L, H](x)$ as a function of viewing angle is obtained as shown in Equation (9) in Appendix.

In data analysis, $P[L, H](x)$ is integrated for

all tested viewing angles, $\Omega = [2.43^\circ, 10.58^\circ]$, as below:

$$T[L, H] = \int_{\Omega} P[L, H](x) dx. \quad (5)$$

A repeated measure ANOVA was conducted to determine whether there were statistically significant differences in $T[L, H]$ over the types of hybrid image. There were no outliers and the data was normally distributed for each group, as assessed by boxplot and Shapiro-Wilk tests ($p > 0.05$), respectively. The assumption of sphericity was violated, as assessed by Mauchly's test of sphericity, $\chi^2(5) = 17.049, p = 0.005$. Therefore, a Greenhouse-Geisser correction was applied ($\epsilon = 0.663$). $T[L, H]$ elicited statistically significant changes over different synthesis methods, $F(1.988, 25.841) = 102.757, p < 0.0001$.

Post-hoc analysis with a Bonferroni adjustment revealed that $T[L, H]$ was significantly different when comparing **Oliva** with **Konishi** (1.820(95% CI, 1.183 to 2.457), $p < 0.0001$), **Oliva** with **Noise** (2.366(95% CI, 1.849 to 2.883), $p < 0.0001$) and **Oliva** with **Color** (2.335(95% CI, 1.922 to 2.748), $p < 0.0001$). Also, $T[L, H]$ was statistically significantly different when comparing **Konishi** with **Noise** (0.546(95%CI, 0.106 to 0.987), $p = 0.012$) and **Konishi** with **Color** (0.515(95%CI, -0.038 to 1.069), $p = 0.076$). However, no statistically significant difference was observed when comparing within our proposed methods, **Noise** vs. **Color**, (-0.031(95% CI, -0.272 to 0.210), $p = 1.0$). The comparison is illustrated in Figure 4.

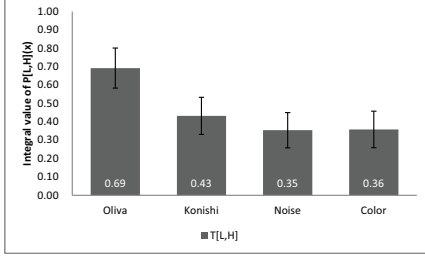


Figure 4: The comparison of average $T[L, H]$ for all observers.

By comparing $T[L, H]$ of the fitted functions, we found that the probability function of **Oliva** has the largest value than others, while our proposed method, **Color**, has the smallest value. This data implied that a hybrid image synthesized by **Oliva** fails to separate spatial frequencies for a wider viewing distance when compared to other methods. It also tells us that $P[L, H](x)$ of our proposed synthesis methods have values only around the middle viewing distance and they are less than those of **Oliva** and **Konishi**. From the experiment result and statistical evidence, our proposed synthesis methods perform better in the separation of spatial frequency images when compared to previous methods.

5.2 Image recognition performance

$P[L, \tilde{H}](x)$ and $P[\tilde{L}, H](x)$ were used to measure image recognition performance. Although the previous section proved that our proposed synthesis methods successfully separated spatial frequencies at most viewing distances, it did not indicate which spatial frequencies were mainly perceived at a specific distance. A hybrid image was created so that the high frequency image was best viewed closely, whereas the low frequency image was best observed from far away. This section explores $P[L, \tilde{H}](x)$ and $P[\tilde{L}, H](x)$ for each image type.

Although the experiment is carefully prepared to reflect as little individual personality as possible, several trends can be spotted from the re-

sults. Even though participants had normal or corrected-to-normal vision, it is still possible that some had better visual acuity than others. Apart from visual acuity, visual attention also varies individually. To deal with such problem, we re-align $P[L, \tilde{H}](x)$ and $P[\tilde{L}, H](x)$ individually by centering all functions to the viewing angle where functions $P[L, \tilde{H}](x)$ and $P[\tilde{L}, H](x)$ intersect. Then, **near** and **far** is set by incrementing and decrementing a proportional angle accordingly. All of the comparisons use an ANOVA repeated measure.

Near viewing distance. The comparison of $P[L, H](\text{near})$, $P[\tilde{L}, H](\text{near})$ and $P[L, \tilde{H}](\text{near})$ across image types is shown in Figure 5. In the comparison of $P[L, \tilde{H}](\text{near})$, we found that there were no outliers and the data was normally distributed for each group, as assessed by box-plot and Shapiro-Wilk tests ($p > 0.05$), respectively. The assumption of sphericity was violated, as assessed by Mauchly’s test of sphericity, $\chi^2(5) = 11.39768, p < 0.0001$. Therefore, a Greenhouse-Geisser correction was applied ($\epsilon = 0.629$). $P[\tilde{L}, H](\text{near})$ elicited statistically significant changes over different hybrid image synthesis methods, $F(1.886, 24.516) = 57.506, p < 0.0001$.

Post-hoc analysis with a Bonferroni adjustment revealed that $P[\tilde{L}, H](\text{near})$ was significantly different when comparing **Oliva** with **Noise** ($-24.582(95\%CI, -33.55$ to $-15.614), p < 0.0001$), and **Oliva** with **Color** ($-32.21(95\%CI, -42.785$ to $-21.636), p < 0.0001$). Also, $P[\tilde{L}, H](\text{near})$ was significantly different when comparing **Konishi** with **Noise** ($-29.834(95\%CI, -39.687$ to $-19.981), p < 0.0001$), and **Konishi** with **Color** ($-37.463(95\%CI, -48.498$ to $-26.427), p < 0.0001$). We also found that our color-inserted method achieved the most $P[\tilde{L}, H](\text{near})$. It was significantly different when comparing **Noise** with **Color** ($-7.628(95\%CI, -12.761$ to $-2.496), p < 0.0001$).

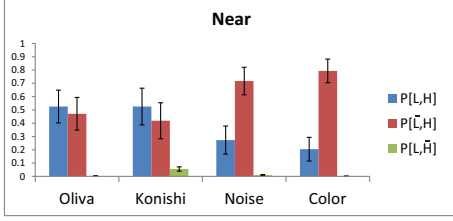


Figure 5: The comparison of $P[L,H](\text{near})$, $P[\bar{L},H](\text{near})$ and $P[L,\bar{H}](\text{near})$ between image types.

From the experiment, we found that at a near viewing distance where most observers perceive high spatial frequency more than low spatial frequency, $P[\bar{L},H](\text{near})$ was the highest for color-inserted method among all methods. In other words, our color-inserted method successfully made the high frequency image more noticeable when compared to other types of hybrid images.

Far viewing distance. At a far viewing distance, we expected that all image types would yield similar $P[L,\bar{H}](\text{far})$ according to human visual limiting acuity, because it is common to perceive only the low spatial frequency from far away. However, we found that this could not apply to a hybrid image synthesized by **Oliva**. The comparison of $P[L,H](\text{far})$, $P[\bar{L},H](\text{far})$ and $P[L,\bar{H}](\text{far})$ across image types is shown in Figure 6.

$P[L,\bar{H}](\text{far})$ was compared across image types by an ANOVA repeated measure similar to all previous comparisons. There were no outliers and the data was normally distributed for each group, as assessed by boxplot and Shapiro-Wilk tests ($p > 0.05$), respectively. The assumption of sphericity was not violated, as assessed by Mauchly’s test of sphericity, $\chi^2(5) = 40.082, p < 0.0001$. Therefore, a Greenhouse-Geisser correction was applied ($\varepsilon = 0.399$). $P[L,\bar{H}](\text{far})$ elicited statistically significant changes over different hybrid image synthesis methods, $F(1.197, 14.359) = 38.92, p <$

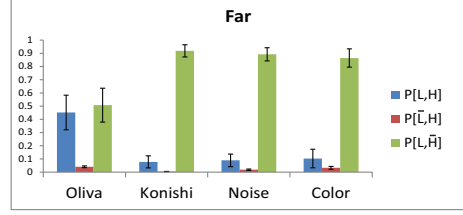


Figure 6: The comparison of $P[L,H](\text{far})$, $P[\bar{L},H](\text{far})$ and $P[L,\bar{H}](\text{far})$ between image types.

0.0001.

Pairwise comparison was done by Bonferroni adjustment and we found that $P[L,\bar{H}](\text{far})$ of hybrid images by **Oliva** scored the lowest, with statistically significant differences when comparing **Oliva** with **Konishi** ($-41.059(95\%CI, -60.414 \text{ to } -21.705)$), **Oliva** with **Noise** ($-38.450(95\%CI, -57.854 \text{ to } -19.045)$, $p < 0.0001$), and **Oliva** with **Color** ($-35.659(95\%CI, -53.154 \text{ to } -18.163)$, $p < 0.0001$). Nevertheless, all of the other pairwise comparison did not show any significant difference, $p > 0.05$.

In addition, the low $P[L,\bar{H}](\text{far})$ observed with **Oliva** could be explained by high $P[L,H](\text{far})$; i.e., both spatial frequencies were visible even at a far viewing distance. An ANOVA repeated measure showed that $P[L,H](\text{far})$ was statistically significantly different across image types ($F(1.20, 14.42) = 33.28, p < 0.0001$ with Greenhouse-Geisser correction, $\varepsilon = 0.4$, because the assumption of sphericity was violated, $\chi^2(5) = 38.81, p < 0.0001$). Pairwise comparison was done by Bonferroni adjustment and we found that $P[L,H](\text{far})$ of hybrid images by **Oliva** scored the highest with statistically significant differences when comparing **Oliva** with **Konishi** ($37.43(95\%CI, 17.76 \text{ to } 57.1)$, $p = 0.0003$), **Oliva** with **Noise** ($36.26(95\%CI, 16.73 \text{ to } 55.78)$, $p = 0.0004$), **Oliva** with **Color** ($34.88(95\%CI, 16.99 \text{ to } 52.77)$, $p = 0.0003$). This shows us that **Oliva** fails to synthesize a

hybrid image from different image structures because the resulting image cannot separate the spatial frequency.

5.3 Discussion

From the experiment, we conclude that our proposed methods in [10] were successful in separating the spatial frequencies at all viewing distances by looking at the integral, $T[L, H]$ of the fitted functions. Hybrid images synthesized by previous methods [3, 6] fail to preserve hybrid image property.

At a relatively close viewing distance, the probability of seeing only the high spatial frequency image, $P[L, H](\text{near})$, was found to be relatively high for our proposed methods when compared to others. Most notably, our color-inserted method yielded the highest probability. This reveals that adding complementary color in the background of the high frequency image strengthens the perception of the high frequency image as anticipated. As $P[L, H](\text{far})$ shows, perception of the low-frequency image is not interfered with when the image is viewed from a proportionately far distance.

With our proposed assessment method, we are able to extract probabilities of seeing both images, or seeing a single image at the same time, while considering individual trial style. It can be adapted to assess other type of ambiguous images, or visual illusions.

Although the experiment was carefully planned so that most issues are covered, we found that some points could be improved. The results suffered from difference in observers' visual acuity. Although every observer had normal or corrected-to-normal vision, some people may have had sharper eyes than others. Viewing distances could be adjusted to suit each observer's acuity. In addition, our proposed analysis method does not extend to use with stimuli like natural images, i.e., hybrids of photographs. Using natural images as hybrid image stimuli in the experiment is less practical and more prone to bias. However, the applications of hybrid image

mostly involve the use of natural images, so it is worthwhile to conduct the experiment using natural images as stimuli. Also, reaction time and measure of accuracy may be considered as measurement methods for recognition performance in future experiments as well.

6. CONCLUSION

This paper presents an assessment procedure for edge alignment-free hybrid images. To assess whether a synthesis method successfully creates a hybrid image that preserve hybrid image property, hybrid images by different methods were compared in terms of the separation of spatial frequency. We introduced four significant probabilities to measure the separation of spatial frequency, along with a suitable experimental procedure. The proposed assessment procedure was tested in an experiment and it revealed that our hybrid synthesis methods [10] successfully achieved the separation of spatial frequency, when comparing to previous synthesis methods.

ACKNOWLEDGMENTS

This work was partially supported by JSPS Grant-in-Aid for Scientific Research (B) (22300030). We specially thank Dr. Haruaki Fukuda (The University of Tokyo) for his fruitful advices.

REFERENCES

- [1] T. Brady and A. Oliva. Spatial frequency integration during active perception: Perceptual hysteresis when an object recedes. *Frontiers in Psychology*, 3(462), 2012.
- [2] N. P. Frederick A. A. Kingdom. Palamedes: Matlab routines for analyzing psychophysical data. <http://www.palamedestoolbox.org>, 2009.
- [3] M. Konishi and Y. Yamaguchi. Hybrid images by local frequency analysis. *Proceedings of Visual Computing, Graphics and CAD Symposium*, 2007. (in Japanese).
- [4] A. Oliva and P. G. Schyns. Mandatory scale perception promotes flexible scene category

rizations. *Proceedings of the 17th Annual Conference of the Cognitive Science Society*, pages 159–163, 1995.

- [5] A. Oliva and P. G. Schyns. Coarse blobs or fine edges? evidence that information diagnosticity changes the perception of complex visual stimuli. *Cognitive Psychology*, 34(1): 72–107, 1997.
- [6] A. Oliva, A. Torralba, and P. G. Schyns. Hybrid images. *ACM Trans. Graph.*, 25: 527–532, July 2006.
- [7] P. G. Schyns and A. Oliva. From blobs to boundary edges: Evidence for time- and spatial-scale-dependent scene recognition. *Psychological Science*, 5(4): 195–200, 1994.
- [8] P. G. Schyns and A. Oliva. Dr. angry and mr. smile: when categorization flexibly modifies the perception of faces in rapid visual presentations. *Cognition*, 69(3): 243–265, 1999.
- [9] P. Sripian and Y. Yamaguchi. Toward a shape-free hybrid image: experimental on visual contrast sensitivity. *Proc. of the 14th Intl. Conf. on Geometry and Graph. (DVD)*, (214): 1–9, 2010.
- [10] P. Sripian and Y. Yamaguchi. Shape-free hybrid image. In *Proceedings of the Symposium on Non-Photorealistic Animation and Rendering*, NPAR '12, pages 11–19. Eurographics Association, Aire-la-Ville, Switzerland, Switzerland, 2012.
- [11] K. Takahashi, M. Hasegawa, Y. Tanaka, and S. Kato. A structural similarity assessment for generating hybrid images. In *Signals, Systems and Computers (ASILOMAR), 2011 Conference Record of the Forty Fifth Asilomar Conference on*, pages 240–243. Nov 2011.

ABOUT THE AUTHORS

1. Peeraya Sripian, Dr., is a lecturer at the department of Media Technology, King Mongkut's University of Technology Thonburi. Her research interest includes visual illusion, human visual system, perception psychology, and image processing. She is a member of ACM. She received a Ph.D. in multidisciplinary science from the University of Tokyo in 2013. Her contact e-mail address is peeraya.sri@mail.kmutt.ac.th.
2. Yasushi Yamaguchi, Dr.(Eng.), is a professor at the University of Tokyo. His research interests lie in image processing, computer graphics and visual illusion, including visual cryptography, computer aided geometric design, volume visualization and painterly rendering. He has been serving as a president of JSGS and a vice president of ISGG. He is also a member of ACM SIGGRAPH, IEEE Computer Society, and SIAM. He received a BE in precision machinery engineering and a Dr. Eng. in information engineering from the University of Tokyo in 1983 and 1988, respectively. He was an assistant professor of Tokyo Denki University from 1989 to 1993. His e-mail address is yama@graco.c.u-tokyo.ac.jp.

APPENDIX

The calculation of probabilities

We assume that all probabilities of seeing spatial frequencies (Equation (1)), trial rates (Equation (2)) and guess rates (Equation (3)) correlate with the probabilities of the answered patterns (Table 1). Equations below show the relationship among parameters.

$$\begin{aligned}
 P[\text{LH}] &= P[\text{L}, \text{H}] + P[\text{L}, \bar{\text{H}}]t[\text{H}, 1]g[\text{H}, \text{H}] + P[\bar{\text{L}}, \text{H}]t[\text{L}, 1]g[\text{L}, \text{L}] + P[\bar{\text{L}}, \bar{\text{H}}]t[\text{LH}, 2]g[\text{LH}, \text{LH}]; \\
 P[\text{LE}] &= P[\text{L}, \bar{\text{H}}]t[\text{H}, 1]g[\text{H}, \text{E}] + P[\bar{\text{L}}, \bar{\text{H}}]t[\text{LH}, 2]g[\text{LH}, \text{LE}]; \\
 P[\text{EH}] &= P[\bar{\text{L}}, \text{H}]t[\text{L}, 1]g[\text{L}, \text{E}] + P[\bar{\text{L}}, \bar{\text{H}}]t[\text{LH}, 2]g[\text{LH}, \text{EH}]; \\
 P[\text{EE}] &= P[\bar{\text{L}}, \bar{\text{H}}]t[\text{LH}, 2]g[\text{LH}, \text{EE}]; \\
 P[\text{LS}] &= P[\text{L}, \bar{\text{H}}]t[\text{H}, 0] + P[\bar{\text{L}}, \bar{\text{H}}]t[\text{LH}, 1]g[\text{LH}, \text{L}]; \\
 P[\text{SH}] &= P[\bar{\text{L}}, \text{H}]t[\text{L}, 0] + P[\bar{\text{L}}, \bar{\text{H}}]t[\text{LH}, 1]g[\text{LH}, \text{H}]; \\
 P[\text{SS}] &= P[\bar{\text{L}}, \bar{\text{H}}]t[\text{LH}, 0]; \quad \text{and} \\
 P[\text{ES}] &= P[\bar{\text{L}}, \bar{\text{H}}]t[\text{LH}, 1]g[\text{LH}, \text{E}].
 \end{aligned} \tag{6}$$

The degree of freedom for answered patterns is 7 ($=8-1$), which is equal to the degree of freedom of unknown parameters, the degree of freedom of the trial rates, four, and the degree of the probabilities of seeing spatial frequencies, three. Therefore, it is possible to solve the equation system, Equations (6). By eliminating $t[X, 0]$ in Equations (6) with Equations (2) and by substituting $g[X, Y]$ with the constant values indicated in Equations (3) and (4), the following equation system in a matrix form is obtained.

$$\begin{pmatrix} P[\text{LE}] \\ P[\text{EH}] \\ P[\text{EE}] \\ P[\text{LS}] \\ P[\text{SH}] \\ P[\text{SS}] \\ P[\text{ES}] \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & \frac{8}{9} & 0 & \frac{8}{45} & 0 \\ 0 & 0 & 0 & 0 & \frac{8}{9} & \frac{8}{45} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{28}{45} & 0 \\ 1 & 0 & 0 & -1 & 0 & 0 & \frac{1}{10} \\ 0 & 1 & 0 & 0 & -1 & 0 & \frac{1}{10} \\ 0 & 0 & 1 & 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{8}{10} \end{pmatrix} \begin{pmatrix} P[\text{L}, \bar{\text{H}}] \\ P[\bar{\text{L}}, \text{H}] \\ P[\bar{\text{L}}, \bar{\text{H}}] \\ P[\text{L}, \bar{\text{H}}]t[\text{H}, 1] \\ P[\bar{\text{L}}, \text{H}]t[\text{L}, 1] \\ P[\bar{\text{L}}, \bar{\text{H}}]t[\text{LH}, 2] \\ P[\bar{\text{L}}, \bar{\text{H}}]t[\text{LH}, 1] \end{pmatrix}. \tag{7}$$

The equation system (7) is solved by multiplying both sides of the equation by the inverse matrix. We can obtain $P[\text{L}, \text{H}]$, $P[\text{L}, \bar{\text{H}}]$, $P[\bar{\text{L}}, \text{H}]$, $P[\bar{\text{L}}, \bar{\text{H}}]$ and $t[X, x]$ under the following constraints:

- $P[\text{L}, \bar{\text{H}}]t[\text{H}, 1]$ must be zero if $P[\text{L}, \bar{\text{H}}]$ is zero;
- $P[\bar{\text{L}}, \text{H}]t[\text{L}, 1]$ must be zero if $P[\bar{\text{L}}, \text{H}]$ is zero; and
- $P[\bar{\text{L}}, \bar{\text{H}}]t[\text{LH}, 2]$ and $P[\bar{\text{L}}, \bar{\text{H}}]t[\text{LH}, 1]$ must both be zero if $P[\bar{\text{L}}, \bar{\text{H}}]$ is zero.

Psychometric function

The cumulative normal distribution function is given as:

$$F(x; \alpha, \beta) = \frac{\beta}{\sqrt{2\pi}} \int_{-\infty}^x \exp\left(-\frac{\beta^2(x-\alpha)^2}{2}\right) dx, \tag{8}$$

where α denotes the location of the function, i.e., threshold, and β denotes the rate of change, i.e., slope.

Function of viewing angle

$P[\text{L}, \text{H}](x)$ as a function of viewing angle is obtained as shown below

$$P[\text{L}, \text{H}](x; \alpha, \beta) = 1 - P[\text{L}, \bar{\text{H}}](x; \alpha, \beta) - P[\bar{\text{L}}, \text{H}](x; \alpha, \beta). \tag{9}$$

AUTOMATED HIGH PRECISION TEXTURING OF 3D-OBJECTS

Martin BORNEMANN, Sebastian MELZER and Daniel LORDICK

Institute of Geometry, Technische Universität Dresden, Germany

ABSTRACT: This paper reports about a 3D-scanning process we established for the purpose of digitizing a collection of mathematical models at the TU Dresden. For this process, we had to match two conflicting requirements: on the one hand we wanted to obtain very precise 3D-models with a highly detailed texture and on the other hand we had to deal with very limited resources. This may be considered as a typical framework for a scientific collection at a university.

Our solution combines two acquisition techniques. For the 3D-model, we use an optical 3D scanner provided by the faculty of mechanical engineering at the TU Dresden[7]. For the textures, we created a setup that automatically produces a series of photographs from different angles with a full frame camera. Both datasets are then merged by means of photogrammetry and additional routines. The latter will be described in detail. In the end we can create 3D objects, which can be distributed and explored with the latest web-technologies.

The whole 3D-scanning process consists of some specially developed algorithms, a chain of open-source modules, and two rather costly modules as well. In the conclusion we will make some suggestions on how to replace the costly modules by open-source components. Thus, we hope to provide a low-cost 3D-scanning scenario, which is useful for other scientific collections - as a whole or in modules.

Keywords: Digitization, 3D-scans, Texturing, Photogrammetry, Collections of Mathematical Models, Virtual Reality

1. INTRODUCTION

While the digitization of printed media is progressing fast, another scientific heritage came only recently into focus for the presentation in online media: the scientific collections at the universities. In 2011, the German Council of Science and Humanities (Wissenschaftsrat) emphasized the importance of the collections as part of the research infrastructure (for further information see [11]). Since then, increasing effort has been made to facilitate the access to the mostly three-dimensional objects by means of online media. Obviously, the precondition is to digitize the objects. One challenge is that the objects can be extremely diverse, even inside a single collection, and another, that the researchers may need to have access to a very detailed representation.

The DAMM-project (Digital Archive of Mathematical Models) at the TU Dresden, funded by the German Research Foundation (DFG), contributes solutions to these challenges and exemplifies them on the basis of a reasonably representative example: the collection of mathematical models at the Institute of Geometry at the TU Dresden. Although mathematical entities are well represented in computersimulations and are rendered on demand in computer-algebra-software, there is close to no digital representation of the historical mathematical models. The models we are dealing with are a cultural heritage, they have a historical dimension, and they have an ongoing effect beyond mathematics. Most of all, they are physical objects we can touch and which therefore create a presence, which can never be

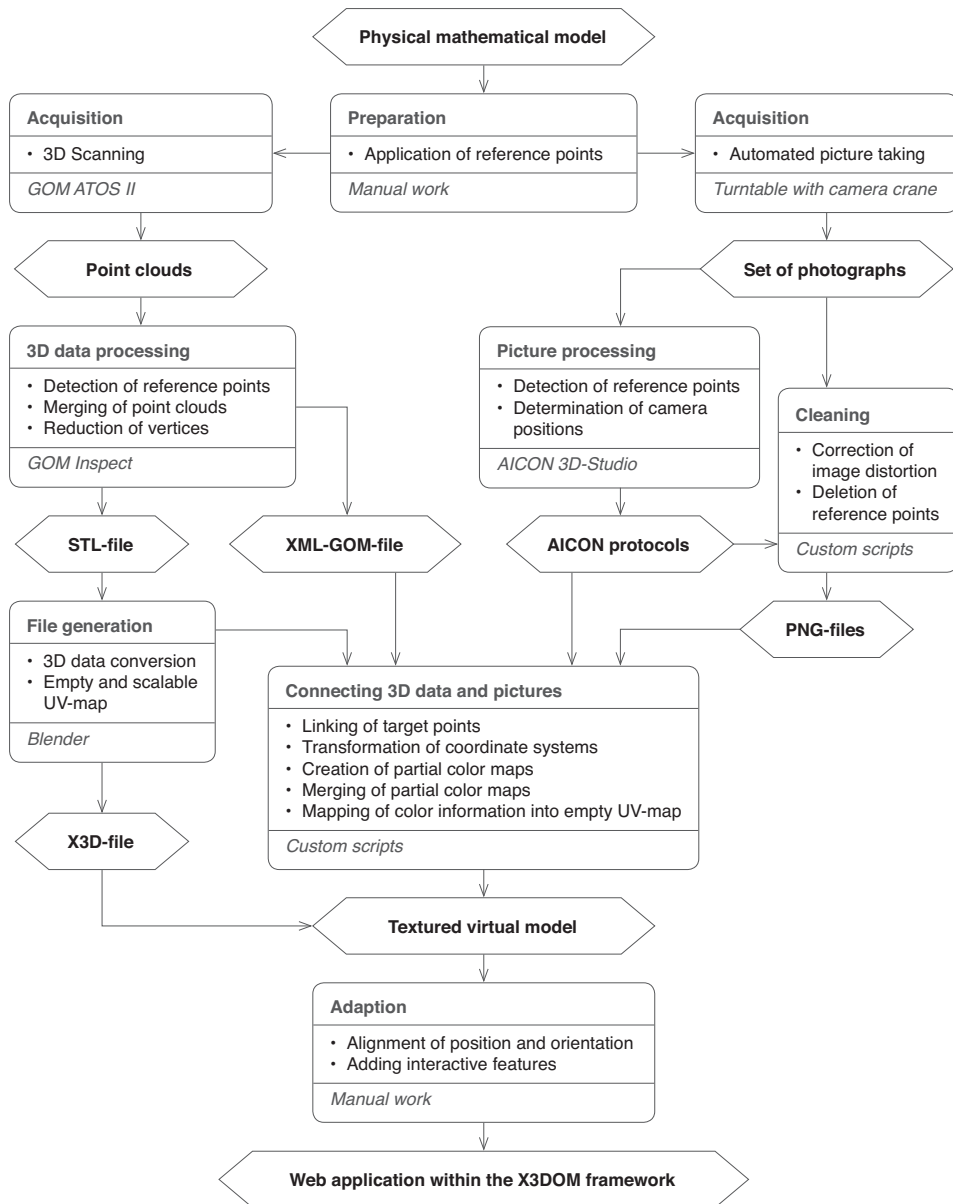


Figure 1: Process of digitization and texturing.

reached by simulations alone. It is quite a task to substitute this presence, spiced with marks, scars and age spots, by online technologies.

The DAMM-project is integrated into the information system University Collections in Germany[4], is linked to the Saxon State and University Library[3], and will contribute to the Europeana[1]. The digitization is paralleled by an indepth scientific exploration and by the development of online technologies like visual assistance in search-functions and cross-media explanations. For the latter we cooperate with the Institute of Software and Multimedia at the TU Dresden and other partners.

In the following, we describe the digitization part of the DAMM-project. Several independent steps are chained together to yield a high-precision virtual model (see figure 1). Those independent steps can be considered as modules, which can be transferred to other digitization projects as well. The whole process is exemplarily illustrated on a surface of constant positive curvature (see [10], XVII, 3(1)).

2. DIGITIZATION PROCESS

2.1 3D data generation

The plain shape of the mathematical model is digitized using a GOM-ATOS-II, an industrial optical 3D-scanner, which provides accurate scans with detailed resolution at high speed[2]. The model is scanned in several takes from all sides. In order to merge the point clouds produced during the single scans, a set of reference points is attached to the object beforehand. These reference points are also a key element in later positioning the texture onto the geometry. The point cloud representing the object is transformed into a triangulated mesh. Then, the file size is reduced by deleting vertices, wherever only slightly curved surfaces have to be represented. At edges, creases and scars the high density of the original point cloud is maintained (see figure 2).

2.2 Automated picture taking

In a second, independent step, a series of pictures of the model is captured automatically by a full frame camera, a Nikon D-600 (see figure 3).

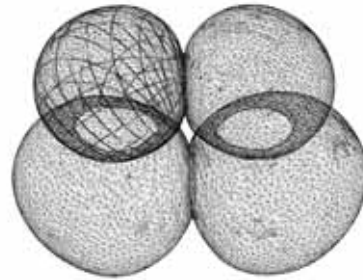


Figure 2: Reduced, triangulated mesh. The triangulation is finer in the upper left part of the model, where there are parameter lines engraved.



Figure 3: Setting for automated picture taking, including model (1), turntable (2), crane with camera (3) and flashlight with diffuser (4).

An ISEL-turntable rotates the model. After each full rotation, a crane moves the camera one step further on a vertical arc around the model. As a result, without further interaction necessary, a set of pictures from equally distributed angles on an imaginary hemisphere is generated, capturing one half of the object (see figure 4). If there is any region of special interest, it is also possible to add single, hand-shot pictures to capture previously hidden details. If other perspectives are needed, the model can be turned around and the whole process can be repeated. We, in general, take a

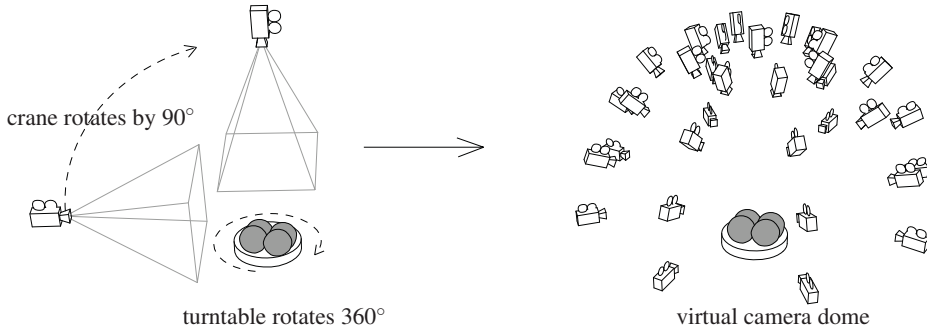


Figure 4: Automated picture taking with one camera, simulating a camera dome

second set of pictures upside down.

Both the turntable and the crane with the attached camera are managed by two ISEL controllers that is accessed via a Python script. Thus, the angle between consecutive pictures can be changed easily. This construction is much cheaper than a multiview dome setup[6] and also allows for the generation of VR-models.

2.3 Determination of the camera positions

For the further processing of the pictures, it is vital to know each relative position of the camera with respect to the object. Unfortunately, this information cannot be retrieved directly from the setup. Even with high-precision industrial turntables, direct measurements result in serious inaccuracies. Therefore, the exact coordinates have to be determined afterwards. The best results are achieved by means of photogrammetry[8]. To support this approach, coded reference points are placed on the turntable beside the model.

These reference points are detected and evaluated by the software Aicon 3D-Studio[3]. Additionally, the software calculates the positions of the initial reference points on the object itself, thus linking the pictures to the scanned geometry. It computes the position of the camera and determines the lens distortion. A Python routine uses this information to remove both the lens distortion and the reference points from the pictures.

At this step of the procedure, there are two sets of reference points, representing two coordinate systems, one of the geometric data and one of the camera positions. At the core of both of these sets are the uncoded reference points placed on the model in the very beginning. In order to unite the geometry and the pictures, a transformation is needed, mapping the reference points of one coordinate system to the reference points of the other. The main problem is that both sets can contain additional points or miss some of the original reference points. While looking only at one set of data, neither the additional nor the missing points can be distinguished from the others. Therefore, all possible combinations are checked and, if there is a certain threshold of matching points reached, numerical optimization is used to find the transformation matrix which maps both point sets best.

2.4 Linking geometry and texture

The next module, the heart of our algorithm, creates the texture for the virtual model on the basis of the pictures. This is achieved in two steps: Getting a partial coloring of the object according to each picture, and then merging the single color maps in a way that finds the true color of each pixel as accurately as possible.

As a preparation, an empty texture map is built by the open-source software Blender, automat-

ically triggered by a Python routine. The generated file contains the 3D-information of the object as well as an empty UV-map, unrolling the triangles and vertices into the plane. By link-

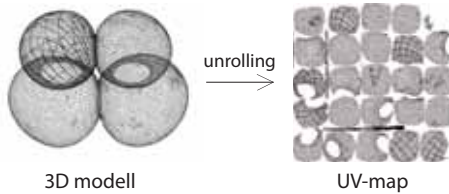


Figure 5: Unrolling geometry into the plane.

ing this file to an image file containing a colored version of the UV-map, the virtual model is finished. Please note that the linked image file can be scaled to any desired resolution, discretizing the object into a finite number of object points, which, in general, will not coincide with the vertices of the triangulated mesh. This means that geometry and texture are strictly separated. The textural information is equally distributed, regardless of the distribution of the vertices. Thus, there can be a lot of textural information in parts, where only a few vertices exist.

2.5 Creating the partial color maps

For every picture taken before, two copies of the empty UV-map are created. One of them is filled with the color information as seen from the current image (see figure 6). The other one stores information about how well every pixel is visible from the current camera position (see figure 7).

A natural approach to retrieving this information would be to use raytracing for every object point, determining all the intersections with the object and coloring only the one which is closest to the camera. Unfortunately, this process is extremely timeconsuming. Therefore, we used the reversed strategy. Instead of intersecting a ray with every triangle of the object, only the intersection with the sensor plane is calculated. As our objects have at least 30.000 triangles and we usually work with more than 10.000.000 ob-

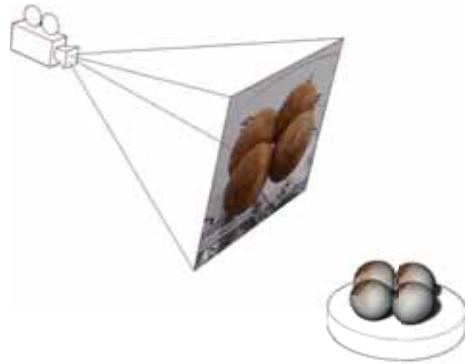


Figure 6: Getting a partial coloring for each picture.

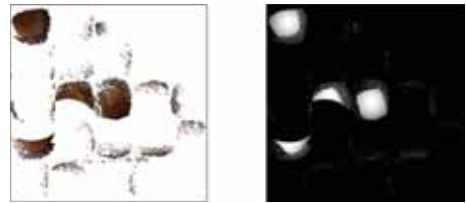


Figure 7: Resulting UV-maps with color and inclination information.

ject points, this is a huge performance difference. Now, three things are stored with every hit pixel of the sensor plane: The object point the ray was sent from, its distance from the sensor plane, and the inclination angle between the ray and the object. If the pixel was already hit by another ray, both distances are compared and only the information with the smaller distance is kept, as the according point obviously hides the other one.

After repeating this process for all object points, the color information of the pixels is transferred to the UV-map, producing a partial coloring of the object corresponding to the current picture (see figure 6). In addition, a second UV-map is filled with the inclination angles of the visible points. Later, this information will be used as a measurement of color fastness. Repeating this process for all the pictures finishes this step.

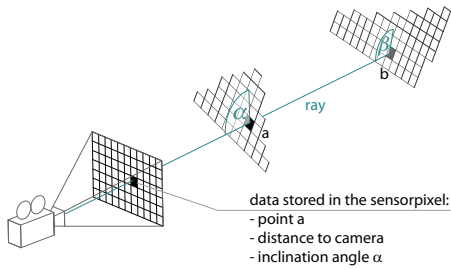


Figure 8: Getting textural information from the current picture.

The careful reader may have noticed that there are still unmentioned problems with determining the visibility. As both the picture and the object are divided into a finite number of points, there is generally no exact matching of hidden points and their obstructions. The first problem arises, roughly speaking, when the resolution of the picture is too high, compared to the resolution of the UV-map.

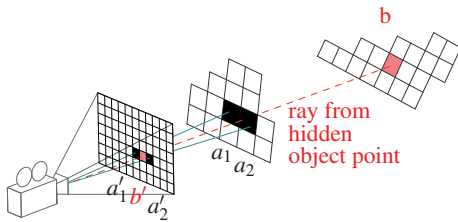


Figure 9: Problem 1: some non-visible object points might be colored.

In Figure 9, object point b is clearly not visible, as it is occluded by the points a_1 and a_2 . Still, the rays through a_1 , a_2 and b hit three different pixels in the sensor plane. To prevent b from being colored, not only its corresponding image pixel, but also its neighbors (up to a case-specific distance) have to be checked for closer points. The second problem is the reversed case of a low resolution in the picture, but a high resolution in the UV-map.

Ray 1 and ray 2 coincide in one pixel on the sensor plane

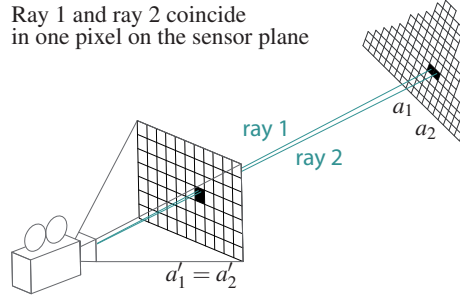


Figure 10: Problem 2: some visible object points might not be colored.

In figure 10, the points a_1 and a_2 are neighbors on the object, but their respective rays hit the same pixel in the image. Without a correction, a_2 would not be colored. The solution to this is introducing a minimum distance for points that are hidden behind each other.

2.6 Merging the partial color maps

In the previous step, for each camera position, one UV-map with color information and one UV-map with the corresponding inclination angles have been created. Now, these maps are merged to produce the final, completely colored, virtual model. For each point on the object, the final color is determined by taking the weighted average of the UV-maps. The weights are calculated using the inclination angles, with perpendicularity resulting in the highest value and not visible points or triangles parallel to the viewing ray having weight 0. Depending on the chosen weight-distribution-function, different textures can be produced.

The easiest function would be to only take the color corresponding to the best inclination angle for each point. This is also the most robust way to do it, as mixing the color information from the different pictures requires a very high accuracy in the steps before. On the other hand, this would also create visible borders on the surface, where the calculation switches from one camera to another. This is due to the fact that the effective

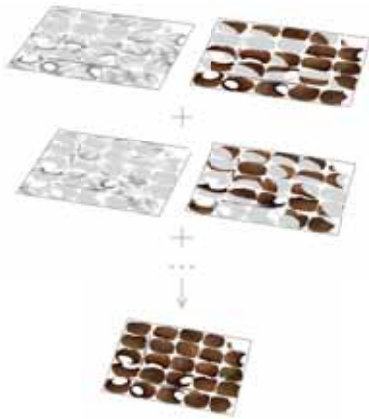


Figure 11: Merging UV-maps.

illumination of the model changes with each picture, so the border region on the object might be darker in one image than it is in the other.

The main advantage of the weighted mixture of the color information is the reduction of shadows and reflections. As the flash light is not directly positioned on the camera, overly bright sections of the object will never be perpendicular to the camera. Therefore, there will always be a non-reflecting point with a higher weight. Even in a different setup, which requires a flash light on the camera itself, this feature is preserved by moving the best weight slightly away from 90° . Similarly, shadows are reduced. With an excessively complicated object, though, they cannot be evaded completely.

2.7 Resulting data

The resulting 3D model(see figure 12) is stored in an X3D file. This format offers a flexible and powerful integration into web-services. With a simple inline-command, the file can be linked to an html-document with an x3dom framework, a JavaScript based interface for X3D [5]. Utilizing the whole spectrum of WebGL technologies, it runs in any browser that supports html 5 and without the need of installing any additional plug-

ins. Several hardware-accelerated shaders are available to present the model with various light-settings. Furthermore, all display parameters can be changed on the fly. This allows, for example, creating dynamic sections of the models by simply changing the z-plane of the x3d-camera. This is definitely a feature that cannot be obtained by working with the physical models.



Figure 12: Final textured model with 64 merged pictures, 6 million object points and 30.000 triangles. In this figure, the reference points have not been removed. This is an easy way of checking the accuracy: if all pictures have been oriented and merged correctly, the reference points will appear sharp and clear.

3. A GLIMPSE OF THE POSSIBILITIES

Having digital versions of the beautiful, historical models allows for a whole new experience of their mathematics. Especially the mixture of information, looking simultaneously at the physical and virtual model, can create an amazingly new understanding, inspiring not only mathematical experts. The next pictures show some of the possibilities, exemplified on a gypsum model of a parabolic cyclide (see [10], X, 5).



Figure 13: Left: Various light settings and shader are available to create a realistic impression of the model. Right: Calculated, purely virtual models and digitized historical ones can be combined symbiotically.

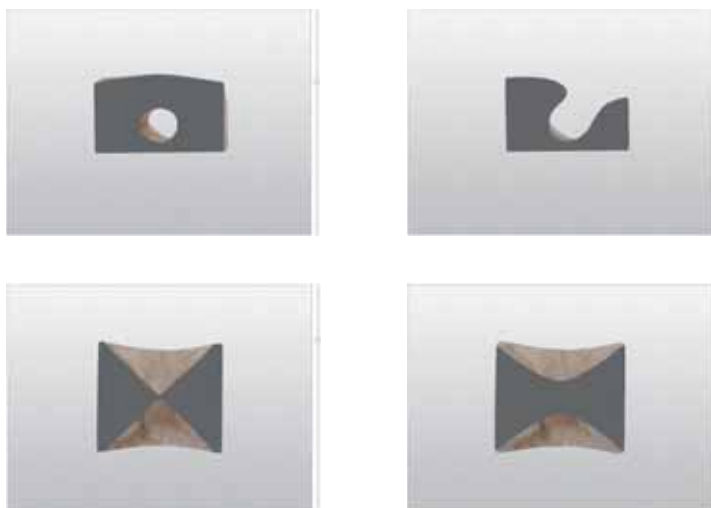


Figure 14: Arbitrary sections of the model can reveal otherwise hidden features. For example, the shown cyclide contains a series of hyperbolas (bottom right), when cut horizontally. Adjusting the height of the cut, they turn into a pair of crossing lines (their asymptotes, bottom left), before going back to being hyperbolas, but rotated by 90° .

4. CONCLUSIONS

When 3D-scanning became popular, many experts expected that in near future, we will have access to more precise equipment at a much lower price. It turned out that today there is a big range of solutions at a low cost, but with a rather low accuracy as well[9]. For this reason, recent attempts concentrated on the problem of calculating better results from poor data. For the collection of mathematical models, we wanted good and reliable 3D-data from the beginning. Luckily, we were able to use an accurate 3D-scanning system. The main issue was to gain textures at a corresponding quality. The digitization process we established offers several, mostly independent modules, which can guarantee this quality and work very well for our situation. But there are still things that can be improved:

At the moment, the 3D-scanning and the picture taking are strictly, even locally, separated. While this does bring considerable algorithmic advantages, it does also mean that a lot of user interaction is needed. It would be preferable to find one low cost solution, which combines 3D-data generation and picture taking, so that the objects have to be put into the procedure only once. This should be automated like our current mechanism and still as accurate as we got used to. One alternative to the commercial GOM-System we use could be VisualSFM[6]. This software produces dense point clouds from a number of photographs. With only a little addition to our algorithm, these point clouds could be used to create a triangulated mesh. Depending on the specific use case, the loss of accuracy, compared to GOM, could be partly compensated by our strict separation of geometry and texture.

The DAMM-project wants to provide routines that other university collections can reconstruct and use, even if they have short resources. Therefore, we want to substitute the two expensive positions in our working chain with open source solutions. As discussed above, the GOM-scanner might be replaced by VisualSFM. In addition, VisualSFM also calculates the camera positions and

the parameters of lens distortion. Therefore, it might also replace the AICON part of our procedure. It remains to be seen if the high accuracy can reliably be maintained in such a setting.

The DAMM-project is carried out by an interdisciplinary research group, which involves architects, mathematicians, and media informatics. Since the collection of mathematical models shall open up not only for mathematicians, but for a wider range of educational, cultural, and entertaining purposes, the mixture of the research group is very helpful and can be recommended for similar projects.

ACKNOWLEDGMENTS

The DAMM-project is funded by the German Research Foundation (DFG) and executed in cooperation with Prof. Dr. Rainer Groh and Hauke Menges from the faculty of informatics at the TU Dresden. It is supported by many colleagues. We especially want to thank Dr. Christine Schöne for her support and the provision of the technology, as well as Prof. Dr. Karin Richter for providing models of the marvelous collection of Martin-Luther University Halle-Wittenberg.

REFERENCES

- [1] Europeana, June 2014. URL <http://www.europeana.eu/>. Accessed: June 2014.
- [2] GOM Optical Measuring Techniques, June 2014. URL <http://www.gom.com>. Accessed: June 2014.
- [3] Sächsische universitätsbibliothek, June 2014. URL <http://www.slub-dresden.de/en/home/>. Accessed: June 2014.
- [4] Universitätsammlungen in Deutschland, 2014. URL <http://www.universitaetssammlungen.de/>. Accessed: June 2014.
- [5] x3dom, June 2014. URL <http://www.x3dom.org/>. Accessed: June 2014.

- [6] C. Schwartz, R. Klein. Acquisition and Presentation of Virtual Surrogates for Cultural Heritage Artefacts. *EVA Berlin*, pages 50–57, 2012.
 - [7] Dr. -Ing. Habil. Christine Schöne, 2014. URL http://tu-dresden.de/die_tu_dresden/fakultaeten/fakultaet_maschinenwesen/imm/ktc/mitarbeiter/schoene. Accessed: June 2014.
 - [8] T. Luhmann. *Nahbereichsphotogrammetrie/ Grundlagen, Methoden und Anwendungen*. Wichmann, Berlin, 2012.
 - [9] P. Kersten, M. Lindstaedt. Generierung von 3D-Punktwolken durch kamerabasierte low-cost Systeme - Workflow und praktische Beispiele. *Terrestrisches Laserscanning 2012*, pages 25–46, 2012.
 - [10] M. Schilling. *Catalog mathematischer Modelle*. Martin Schilling Verlag, Leipzig, 1911.
 - [11] Wissenschaftsrat. Empfehlungen zu wissenschaftlichen Sammlungen als Forschungsinfrastrukturen, 2011. URL <http://www.wissenschaftsrat.de/download/archiv/10464-11.pdf>. Accessed: June 2014.
- Dresden University of Technology, Germany. His research interests are combinatorial triangulations, embeddability of low genus surfaces and computer aided geometric design. He can be reached by e-mail: sebastian.melzer@tu-dresden.de or through postal address: TU Dresden, Institute of Geometry, 01062 Dresden, Germany.
- 3. Daniel Lordick studied architecture at the TU Berlin and the Carleton University in Ottawa. He is professor for Geometric Modeling and Visualization at the TU Dresden, Institute of Geometry. He can be reached by e-mail: daniel.lordick@tu-dresden.de.

ABOUT THE AUTHORS

1. Jens-Martin W. Bornemann is a freelance architect. He also works as a research assistant for the DAMM-project. His research interests are generative design, computational architecture and CAD-visualization. He can be reached at www.bornemann-architektur.de or by e-mail: jens-martin.bornemann@tu-dresden.de
2. Sebastian Melzer, Dipl.-Math., is a PhD student at the Institute of Geometry at

AUTOMATIC RANGE SCAN POINT CLOUD REGISTRATION USING HIERARCHICAL LEVELS AND FEATURE RECOGNITION FILTERS

Alejandro ALONSO and Min-Hyung CHOI
The University of Colorado Denver

ABSTRACT: Range scan point cloud registration problem has been well studied from the extensive exploration of the range images' degrees of freedom to the most commonly used iterative closest point (ICP) approaches. However most iterative point-wise methods work well only when the range images are close enough to solution configuration, and there's no guarantee that ICP would work or not from a given arbitrary configuration. Furthermore, the pair-wise nature of registration demands that each corresponding range image sequential pair must be known beforehand for a complete all-around registration from a batch of unordered range scan to build a complete 3D structure. This paper proposes a novel approach that aligns a set of unordered range images without knowing the correspondence between pairs to the point where ICP is guaranteed to succeed in its finer level registration. The registration task is nicely divided into two phases: the coarse registration which approximates the range images to a probable solution space, and the fine registration that takes the pre-aligned set of point cloud and further refines them to an optimally registered configuration. By incorporating a multilevel hierarchical structure similar to the ones used in bounding volume collision detection methods, this paper aims to reduce unnecessary processes in registration pipeline.

Keywords: Registration, range image, point cloud, laser 3D scanning.

1. INTRODUCTION

In recent years the number of scanning devices available in the market has been multiplied, resulting in greater accuracy and cost competitiveness of 3D scanning. Because of this greater number of scanning devices, varied implementations of 3D reconstruction using these devices have been developed, from recreating sceneries for entertainment purposes and building virtual organs for medical purposes, to terrain reconstruction for robotic navigation. Besides the current applications, the nascent 3D printers technology greatly increases the possible number of future applications for 3D scanners. However, scanning and 3D reconstruction technology is limited to the range of view of the sensory devices. The foundation of the scanner limitations is the necessity of a set of scans that complete the total surface of the object from

different points of view of the object. This set of images of the same object holds an unknown relation between them, a transformation matrix and the problem of finding the transformation matrix between corresponding range images is known as registration.

The problem of registration has been extensively explored with several different successful approaches. The most commonly used registration method is ICP which has a very high success rate for cases in which the range images are close enough to the correct orientation, but a very low rate for objects in orientations far from their original relation; this situation has lead to the divergence of registration approaches, the coarse and fine registrations. Coarse registration makes an initial guess that approximates the range images close enough to their original relation and the fine registration that takes the initial guess and further refines it,

a commonly used method for fine registration is ICP. Another limitation of these approaches is the pairwise nature of them which restricts the possible inputs to pairs of previously known corresponding range images.

This work considers the possible limitations from some previously proposed approaches and addresses registration with a process that is able to coarsely register a set of unordered range images with no information of their orientation or translation to positions close enough to be refined with the final step of ICP. To do this, it uses surface descriptors and multiple hierarchical levels that are able to reduce the influence of external noises inherent to feature recognition algorithms as well as the large number of computations by using feature recognition filters that disregard improbable matching surfaces.

The remainder of this paper is organized as follows: Section 2 discusses some of the previous researches in the registration matter. Section 3 describes the hierarchical structure used to simplify the computations of the registration process. Section 4 explains the simplified computations used for the coarse registration at the pairwise phase and then the proposed metric implemented to complete the coarse registration at the batch phase. Section 5 explores the sets of experiments applied for the verification of the proposal. Finally, Section 6 contains the conclusions of this work and references for possible future works.

2. RELATED WORK

3D point cloud registration is a problem that has been widely studied. However the nature of these works has been mostly dedicated to a pairwise registration excluding information that could be used to identify correspondences from several possible pairs. In [1] Fukai and Xu explore the six degrees of freedom of a pair of point clouds and evaluates them with ICP to find the best transformation between them, however the utilization of the ICP alignment error still finds absolute minimum with incor-

rect matches and even with a pairwise registration the search of all possible combinations is computationally expensive; as a solution to the great number of computations Makadia et al [2] propose a feature recognition solution that is able to find the most representative points in both range images named constellation images; the registration is done considering just these constellation points, even though the successful results of this approach, the introduction of a bumped surface caused by a low quality scanner will create the problem of identifying non-corresponding features for the constellation images; In [3] the exploration of the degrees of freedom is reintroduced but this time the computations are sped up with a random point sampling scheme and an octree to compute distances, the result is improved in the case of computational expenses however the random sampling introduces extra inaccuracy with the ICP error evaluation. [4] proposes the use of a random point sampling and surface descriptor named tensor that identifies the intersection with the other range image and obtains a metric that compares the possible pairs from a complete set of range images, however the fact that it creates the tensor grids from randomly selected points translates into the possibility that there may not exist correspondence between the points being registered. Finally [5] uses a feature description approach named local invariant feature; it describes the neighborhood for a point with a single scalar value then it compares the features and finds the corresponding points, the limitation with this approach is that the scalar descriptor may be identical in many different points for special cases like cubic shapes, where the different planes rotated will have a very similar local invariant feature in each face or even in repetitive shapes which could lead to an incorrect alignment.

3. HIERARCHICAL STRUCTURE

Fine level registration has a very efficient solution in ICP for the cases in which the initial guess is close to the actual transformation. This

situation leads to the necessity of finding a proper coarse level registration. Considering the great amount of points included in each of the scans, the possible solutions search becomes highly expensive, even more considering the fact that a set of range images is to be compared in a round robin to find the most probable correspondences. To solve the point cloud complexity problem it is proposed to simplify the point clouds in an organized structure that allows the acquisition of point vicinity and a hierarchical structure without further computations; this structure is similar to the structures used in collision detection problems. The result of these structures used in collision detection is that they limit the number of elements at each level improving the complexity for the range images correspondence search and by simplifying the point clouds it reduces the induced noise influence during the feature recognition. The rest of this section will be focused in the explanation of the methodology used to create the hierarchical structure for the program.

3.1 Simplified Points (Level 1)

The first level of simplification this work proposes is an ordered, quasi-evenly distributed representation of the point clouds stored in a 30x30 array. This array is able to reduce the point clouds into a much smaller number of points that will be used for the coarse registration process instead of the thousands of points contained in the range images.

The process starts by obtaining an axis aligned bounding box (AABB) for each range image. The AABB dimensions are used to obtain the rough dimensions of the range images from the point of view of the scanner. The maximum and minimum values of the three axes are updated including all the point clouds to be registered. After that these dimensions are used to obtain the local surface size that will be used to compute the simplified points as in the following Equation (1):

$$Size(x, y, z) = \max(\max(x_p, y_p, z_p) - \min(x_p, y_p, z_p)) \quad (1)$$

The next step is to use the local surface dimen-

sions that were obtained to create n local bins that contain the points within the bin boundaries, to do this a binary search tree that returns the points within a certain range is used; the ranges for each bin are obtained by dividing the range image:

$$HB(i, j) = (n(Size(x), n(Size(x) + Size(x))) \quad (2)$$

$$VB(i, j) = (n(Size(y), n(Size(y) + Size(y))) \quad (3)$$

Once all the bins have their corresponding points contained, the following step is to compute AABB for each bin; the obtained AABBs are then used to obtain the simplified points:

$$SP(x, y, z) = \left(\frac{AABB(x_1) + AABB(x_2)}{2}, \frac{AABB(y_1) + AABB(y_2)}{2}, AABB(z_1) \right) \quad (4)$$

The points obtained are stored in the array which maintains the information ordered. The ordered array optimizes the access to the points stored and finds the neighboring point indices with a simple formula instead of computing distances with all the points:

$$NP(i, j) = (i \pm 0 \parallel 1, j \pm 0 \parallel 1) \quad (5)$$

Then for the feature recognition filters, the normal vertex of each simplified point is computed first computing the surface normals of all the neighboring quads:

$$QN(i, j) = \frac{(P_{i+1,j} - P_{i,j}) \times P_{i+1,j+1} - P_{i,j+1}) + (P_{i,j+1} - P_{i+1,j+1} \times P_{i,j} - P_{i,j+1})}{2} \quad (6)$$

Finally, the vertex normals are obtained using the surface normals of the neighboring quads:

$$N_p(i, j) = \frac{\sum_{s=0}^q QN_s}{q} \quad (7)$$

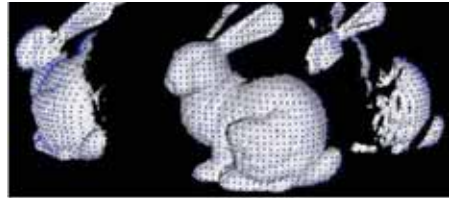


Figure 1: Example of Simplified Points.

3.2 Bounding Spheres (Level 2)

Since the first level of simplifications results still in a great computation expense during the

registration, a second level of simplification is proposed to further reduce it. However, the use of surface description is necessary to filter improbable matches and so far previous methods lack the proper utilization of hierarchical levels and their surface information. To solve this problem, the second level is also used to contain surface descriptors from the previous level. The second level is where all the surface information is kept which will later be used for the matching filters, making it the most frequently accessed level in the hierarchical structure. The structure representing this level is an ordered array of bounding spheres. The bounding spheres were chosen because the sphere intersection determination is the fastest of all the bounding volumes. Each bounding sphere contains a 3x3 level 1 simplified points allowing it to have an even description of the surrounding surface around the center point and an easy computation of the neighboring points which can be automatically found using the level 1 ordered array.

Using the 9 points for each corresponding surface patch, the bounding spheres are computed like this:

$$R_0(i, j) = \max \text{length}(AABB_x, AABB_y, AABB_z) \quad (8)$$

$$\text{Center}(i, j) = \left(\frac{AABB_x}{2}, \frac{AABB_y}{2}, \frac{AABB_z}{2} \right) \quad (9)$$

$$R_f(i, j) = \begin{cases} \text{dist}(p_i, \text{Center}) > R_f \rightarrow R_f = \text{dist}(p_i, \text{Center}) \\ \text{else} \rightarrow R_f = R_f \end{cases} \quad (10)$$

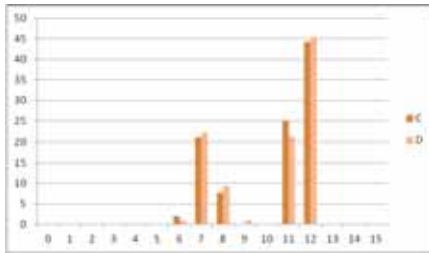


Figure 2: Example of a Local Feature Histogram.

Then using the technique of surface patch

description proposed in [8], with all the possible pairs of points contained within each bounding sphere, the local feature histogram is created. For each element of the Level 2 structure a histogram is stored so that it can be compared to filter improbable surface matches.



Figure 3: Example of Bounding Spheres simplification.

3.3 Bounding Boxes (Level 3)

While executing registration at the second level, it can be observed that several computations still happen even where the collision between range images is not happening. Considering how collision detection algorithms solve this problem with higher hierarchical levels that simplify the mesh to avoid unnecessary collision computations, the third level encompasses level 2 bodies to reduce computations. The third contains a 2x2 bounding spheres set within an ordered array of AABBs. The choice of bounding volume in this case is because not all the resulting simplified points meshes are approximately even in their x and y dimensions and that causes that the resulting bounding spheres become too large in order to contain the points from level 2 which would not be an accurate representation of the range image. It was found in our experiments that a correct surface match was volatile due to the level of detail lost at this simplification scheme; because of this, the third level is used solely with the purpose of collision detection.

After this level, further simplification of the point cloud becomes so rough that it does not represent the range image and as a result it only increases the number of operations needed to find if the range images are overlapping.

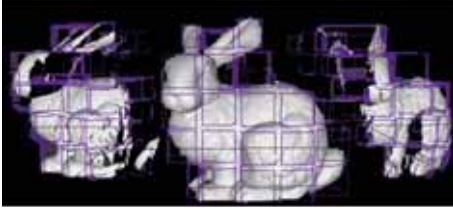


Figure 4: Example of AABBs simplification.

4. COARSE REGISTRATION

One of the main purposes of this approach is to find an accurate initial guess that is close enough to the actual transformation of a set of range images so that the ICP algorithm can further refine the transformation using the point to point distance optimization. The related works that have been mentioned so far in this work cover a wide area of ideas from brute force approaches where the degrees of freedom are exhaustively explored to the feature detection algorithms that refine the meshes to a small number of points that are later optimized with ICP variations.

Our proposal to solve this problem is divided in two stages: first the pairwise registration in which all the possible pairs of range images are explored to find possible matching surfaces, and second the batch registration that decides the order in which range images should be registered. In order to find the batch registration it is necessary to obtain an evaluation metric from the pairwise registration process so that it can be identified which is the best combination. The proposal in this case is to take advantage of the hierarchical structure to speed up the computation and use the feature descriptors to reduce even further the exploration.

4.1 Pairwise Registration

The main idea of this work for the coarse registration integrates aspects of the related works that have been explored in Section 2. The exhaustive exploration of the combination of points is used to find the orientation and translation of the range image that is being registered. This approach takes elements proposed

in [1] and [3] where they explore exhaustively the transformation space; but our solution proposes to limit the range of transformations using a simplification structure. Considering [6], a hierarchical structure was implemented; however, instead of neglecting all the information included in the point clouds, a multi-level hierarchy is used where the level of detail keeps incrementing as it advances in the hierarchical tree. The hierarchy levels serve to improve the process as well by speeding up the verification. Another contribution in this work is the use of feature filters, the proposed algorithm takes the feature description to filter out transformations that are not probable by using just similar surfaces to register pairs of points. This modification maintains all the information available to execute the metrics of the registration quality instead of limiting the number of points available for the registration where some of the remaining points may not be part of the overlapping area between the range images.



Figure 5: Pairwise Coarse Registration Example 1.

The first step is to explore all the possible level 2 pairs formed from the two range images. The process starts by choosing the central point from a bounding sphere from the source range image and the destination range image; in this case the central points are already known thanks to the ordered arrays. Then the histograms corresponding to each bounding sphere are compared as a filter, in the case that the pair meets the criteria, it continues to the next step, otherwise the pair is discarded. For the next step, the normal from that central source point in the range image to be registered is then rotated to obtain the same orientation of the central point in the destination range image and translated to make the points coincidental. Af-

ter that, the orientation needs to go through a correspondence filter; if it fails, the pair is discarded; otherwise a chain of overlapping volumes is searched with the help of the hierarchical structure. All the possible pairs of points from level 2 are searched and the ones that have a value above the threshold are saved for the final metric evaluation. If no pair is found after all the tests then the algorithm determines that there is no correspondence between the range images pair.

The algorithm proposes the use of two filters, a filter for the surface similarity and a filter for the correct orientation. The use of these filters intends to reduce all the possible mismatches as well as the total number of deep explorations in search of a chain of coincidences that would result time consuming and unproductive.

The surface similarity filter uses the feature recognition techniques used in [6]. The surface histogram stores the information of the curvature inside the level 2 bounding volume and it is used to compare the pairs. A matching patch of surface from the source range image should be similar to the matching patch of surface from the destination range image. Using the histograms distribution, they are compared bin to bin and a percentage of similarity is computed, if the percentage is greater than the threshold percent then the points are treated as possible matches and kept for further exploration.

The next filter checks that the simplified points contained in the level 2 bounding spheres have a matching surface orientation between the explored pair. It starts by finding the closest point to each of the level 1 points between the source and destination range images; each of them should have a distance less than the half the length of the level 1 bins ($MD/2$) if not the pair is rejected. Then for each point, the closest point normal vector difference should be less than 10 degrees, if more than 3 of the points fail it is rejected.

And finally, the incremental orientation; the central point in both range images are incre-

mented, if the incremented points are not the closest to each other, then the source is rotated around the normal vector until the incremental closest points are coinciding.

The next step is to find the number of level 2 bodies intersecting each other from both range images, this process is named chain search. In this step all the surviving pairs are explored. For each explored pair the central points of the bounding spheres are aligned first.

The chain search is where the suggested solution takes advantage from the techniques used in collision detection and the constructed hierarchy structure. The chain starts with the current registered level 2 pair and checks for collision between them. The process is repeated to all the level 2 spheres that share the level 3 parent node with the origin. It then follows with the neighboring level 3 volume from the source range image, it finds if it is intersecting with some other level 3 volume from the destination range image. Then it searches the intersections on the level 2 spheres and for the found collisions it adds the pairs to the chain. This process grows at the level 3 volumes and explores deeper only in case of probable chains links inside the explored branch.

After obtaining the chains of coincidences the pairs go into the last filter. The chains of coincidences whose number of elements are greater than the threshold are stored as possible final registrations and evaluated to determine the best possible option.

All the pairs of points that have a chain larger than the threshold are then explored in detail to have its closeness measured. The tool used for this metric is a Quadtree that contains the destination range image. This Quadtree obtains the minimum distances from all the level 1 points in the source mesh to the destination mesh level 1 points in asymptotical time of $O(\log N)$. There are two possible ways of evaluating the registration, using the average error or using the number of coincidences.

The Quadtree is constructed from the root to the leaf nodes; it starts by obtaining the AABB

from all the points in the level 1 array. The AABB is then split by the half along its two greater axes.

All the points are divided in the 4 resulting sections of the parent node. This process is repeated iteratively until all the leaf nodes have less than 8 points. The Quadtree receives the 3D points from the source point cloud and it computes the closest leaf node to each point from the destination point cloud then returning the minimum distances between points.

The registration starts with the point p in the level 1 hierarchy from the source. It computes the minimum distance from the point to the destination with the Quadtree. Then error accumulation is added with the minimum distance and after repeating the process for all the points available in the source range image the error is computed like this:

$$Error = \frac{\sum_{s=1}^n \min Dist(p_s)}{n} \quad (11)$$

The error is compared with all the probable registration chains and the registration with the lowest error is chosen as the correct coarse registration.



Figure 6: Pairwise Coarse Registration Example 2.

Besides this metric, our approach includes a different metric. The other method uses the Quadtree also to compute the minimum distance from each point of the source image to the destination image, but in this case instead of using the distance to compute the error, the value is compared and if it is less than the threshold, an occurrences counter will be incremented. The total number of occurrences is then used as the metric.

4.2 Batch Registration

The solution assumes that the range images re-

ceived contain no information concerning their order or which range image is supposed to be registered with the others. To solve this problem, first it registers all the possible pairs of range images formed with the first range image; the metrics returned from the pairwise registrations are used as reference. Depending on the metric, the minimum error or the maximum number of occurrences, determines which is the corresponding range image.

The process is repeated with the next range image and the remaining available. The Quadtree is created with all the range images that have already been registered to increment the probability of finding the correspondence. These steps are repeated with all the remaining range images available or until the remaining point clouds have been found not coincident with the all the range images.

The result of this approach however is dependent of the point clouds exploration order which could lead to choose an incorrect first alignment that would close the cycle prematurely, because some remaining point cloud had no possible registration with the remaining range images. To avoid this problem, the solution was modified to a comparison between all the possible pairs of point clouds.

The program compares the point clouds in a round robin to find the registration between them, from this possible registration the error is stored as a metric for that match. After all the matches are computed, the point clouds are sorted increasingly according to their number of possible matches. The possible matches are those pairs of point clouds that have at least 25 percent of level 2 collisions. Starting by the point cloud with the least number of possible matches, the algorithm chooses the match that has the smaller level 1 error after the registration and then continues to the next point cloud in the sorted array. With this the algorithm finds a route that goes through all the point clouds with no inner cycles and that has the least error possible. However, the results of this approach were not satisfactory; this could be

explained by the fact that the minimum distance average depends of all the points and even the points outside the overlapping region account for the error, leaving a big space for local minima. To improve the influence of the non-corresponding area, the error metric was substituted by the number of corresponding points between the pair. This approach depends only of the points that are included in the found overlapping area and works as a better judge for the correspondences.

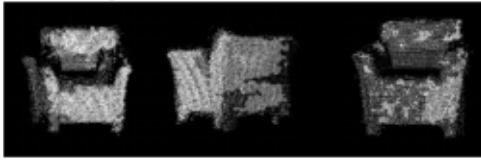


Figure 7: Batch Coarse Registration Example 1.



Figure 8: Batch Coarse Registration Example 2.

5. RESULTS

The experiments implemented to test the registration of sets of range images were developed first to denote the optimal values needed to obtain a better performance of the hierarchical structure, second to understand the characteristics needed to obtain a successful registration and third to find if given the correct range images set the algorithm could perform a correct registration regardless the initial position of the images.

5.1 Parameter Settings

The first set of experiments utilized a batch of range images of both scanned and virtual point clouds to define the simplified points' array size. Since we know that the level 2 structures consists of 3 x 3 level 1 points clus-

ters, then the squared array of level 1 must be a multiple of 3 to obtain an exact division in the previous hierarchical level.

Table 1: Level 1 Array Size Registration.

	Hier- chy (s)	Coarse (s)	Fine (0.85, 200) (s)	Fine (0.50, 0.50) (s)
Chair (18x18)	2.285	2.264(*)	6.855(*)	1.328(*)
Bunny (18x18)	14.850	0.982(*)	70.048(*)	50.359(*)
Chair (24x24)	2.315	6.926(*)	8.219(*)	1.556(*)
Bunny (24x24)	15.303	1.968(*)	71.734(*)	15.361(*)
Chair (30x30)	2.100	26.500	6.953	1.273
Bunny (30x30)	15.225	8.491	246.414	39.116
(*) = unsuccessful registration				

The level 1 array sizes were set and the time of the pre-computations, the time of the coarse alignment and the time of the fine alignment were obtained for the analysis. From the results we can observe that for lower values of the level 1 arrays, the pre-computation time is very close. For the registration we can observe how the time increments exponentially as the array size grows for both types of registration, however the lower sizes return no convergence for the batch registration; the cause of this is that the lower sizes fail to represent the level of detail needed for feature recognition, therefore we can determine that the optimal level 1 size $N=30$. It is also important to remark that the fine registration returns a very similar result after the coarse registration with the configuration of 200 iterations and 85 percent of sampling and the configuration of 50 iterations and 50 percent of sampling. The explanation of this is that the coarse alignment is close enough to the optimal solution of the ICP algorithm that the sampling and the iterations can be reduced and it will still return a close fine registration.

The histogram comparison works as a filter to speed up the registration by eliminating the pairs that are not probable corresponding surface patches because of their different surface properties. This means that a great number of

improbable registrations will be rejected and the time consumed analyzing these incorrect solutions will be saved.

The second experiment took a pair of virtual range images and the percentage of acceptance was incremented starting at 50 percent until 95 percent; this time was also saved for the analysis and the outcome of the registration.

It can be observed that the registration time is reduced dramatically as the parameter is increased; however at the when we reach the 90 percent of similarity no correct registration is found. That is because the representations of the surface patches are very similar to the original but not identical due to the information lost during the simplification process. According to these results we can conclude that an 80 percent of similarity is enough to filter the improbable registrations and to keep a safe margin of detection of surface coincidences.

Table 2: Similarity Percentage Times.

Per-centage	Mesh1 (s)	Mesh2 (s)	Mesh3 (s)	Mesh4 (s)	Mesh5 (s)
0.5	7.625	22.818	16.798	64.602	7.426
0.55	7.305	19.657	16.784	59.149	7.459
0.6	7.215	19.114	16.834	37.205	7.425
0.65	6.894	18.875	16.525	26.584	7.387
0.7	6.562	17.122	13.571	24.355	7.321
0.75	5.682	13.649	11.987	22.795	7.258
0.8	4.868	11.898	10.694	20.378	7.258
0.85	4.126	10.489	9.167	13.488	6.951
0.9	3.628	8.399	8.027	11.315	5.268
0.95	2.865	7.479	8.519	11.537	6.121

The minimum point to point distances are important for the creation of the chains of coincidences and for the metric. The next experiment explored the effects of the minimum distance. Considering that level 1 points are a grid that is spaced at a maximum distance of MD, the experiments take MD as the threshold value and then a percentage of it is used incrementally to observe the outcome of the registration.

The results of the experiments reflect that a $0.50md$ is the optimal threshold value. The explanation for this is that a point from the first point cloud that is correctly registered at the second point cloud could be at maximum in the

Table 3: Similarity Percentage Results.

Percent-age	Mesh1 (s)	Mesh2 (s)	Mesh3 (s)
0.5	OK	OK	OK
0.6	OK	(*)	OK
0.7	OK	(*)	(*)
0.8	(*)	(*)	(*)
0.9	(*)	(*)	(*)
1.0	(*)	(*)	(*)
1.1	(*)	(*)	(*)
(*) = unsuccessful registration			

exact center between two points in the grid. Greater values than this would account for incorrect coincidences and the lower values would fail to account a number of correct coincidences.

5.2 Algorithm Validation

Once the optimal parameter settings were defined, a series of experiments were developed to prove the effectiveness of the solution. The experiments test the overlapping area between the range images, the initial configuration of the range images and compare the proposed algorithm with the ICP alignment.



Figure 9: Overlapping Percentage Alignment.

To determine if a registration was correct, a four step metric was proposed. The test consist of the ICP error before the fine registration and after the fine registration checking if the error has indeed been optimized by the ICP algorithm after the coarse registration. The next step measures the time it took to obtain the complete registration including the hierarchical structure creation. And the final step consists of the intervention of a person; that is because for the computer the registration could be considered as a success if it has found any similar surfaces large enough to pass the threshold, even when it is an incorrect registration. It can

only be determined if it is indeed correct or not by manual inspection. The last metric is a Boolean indicator that describes if a user observes a correct registration.

Table 4: Overlapping Percentage Alignment.

C.Points	%	Alignment
531	91	OK
449	77	OK
412	71	OK
334	57	OK
299	51	OK
223	38	(*)
(*) = unsuccessful registration		

Table 5: 4-Step Metric.

Mesh	Coarse Error	Fine Error	Visual	Time (s)
1	0.00134	0.00027	OK	38.165
2	0.01503	0.00121	OK	26.702
3	0.00414	0.00027	OK	39.254
4	0.02211	0.01199	OK	3.931
5	0.08741	0.04074	OK	4.202

This test was conducted to five pairs of corresponding point clouds. The point clouds were chosen as two pairs of scanned data point clouds that included a moderate level of noise that did not deform the overall shape of the objects (chair and car), then a pair of virtual data point clouds with large uniform surfaces (car), and finally two pairs of high quality real data point clouds with very unique surfaces (armadillo and dragon).

The results of this experiment show that the minimum percentage of overlapping area is near the 50 percent. This value could be reduced if all the threshold values were modified, however that would eliminate all of the filters advantages to reduce the computations proposed in this work.

The next set of experiments were use to determine the reliability of the algorithm depending on the initial orientation of the range images was tested. The two independent experiments check first the impact of the original position in the outcome of the algorithm and the

second experiment checks the impact of the original orientation in the outcome of the algorithm.

The experiments were conducted on a pair of virtual range images. The translation test consists of moving one of the range images along the three axes, and registering the outcome and the registration time. The measuring unit for the distance in this case is MD.

The results of the experiment showed that the distance between point clouds did not affect the outcome of the algorithm, which is an expected result because the registration problem tries to optimize the distance between points in every proposed solution.

The orientation test consisted of rotating one of the range images around the three axes before registering the point clouds. The experiment showed that the registration algorithm works in almost every original condition with the exceptions of the 90 and -90 degrees around the z axis, this is because it does the rotation angle computations as a 2D planes and when a vector is aligned with the axis it would need a division by 0 which is an error. To solve the problem the orientation of that range image can be slightly modified to avoid the division by 0.

6. CONCLUSIONS

The main objective of this work is to develop an algorithm that is capable of registering unordered range images with error induced from low definition scanning devices. To solve this problem, this work proposed an innovative method for the automatic range image registration that combines different types of registration: the optimization of a metric, the feature recognition and an exhaustive search of the transformation range with a multilevel hierarchical structure that reduces the complexity of the match evaluation but maintains details accessible to differentiate possible matches.

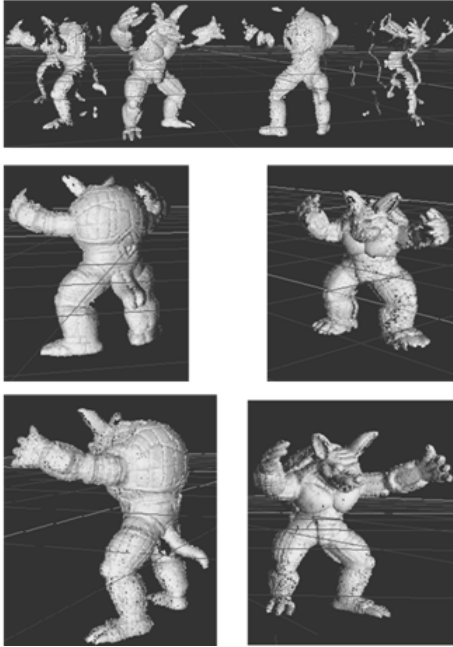


Figure 10: Coarse and Fine Registration Example.

After the experimentation conducted on the algorithm, it can be observed that the proposed hierarchical level approach optimizes the asymptotical time of the registration problem. The utilization of the feature recognition filters also reduces the search space by eliminating all of the improbable solutions.

Compared to other methods with hierarchical levels or feature recognition, instead of consisting only of two levels that are the complete point clouds or a few points, the utilization of the multiple levels is an improvement in the sense that it can keep valuable information for the registration easily accessible and be used as well as a simplification of discarding unviable solutions.

Finally, the definition of the simplified mesh consists of a known maximum number and this fact compared to the approaches referenced in

this work will reduce all the range images independently of the number of points in the range image. This approach applied to low definition scanning devices, like some of the sets used in this project, results in a simplification that is detailed enough to find the surface properties but coarse enough that neglects the error induced by the scanner noise.

The proposed solution is limited to cases in which the objects shape is unique enough to be differentiated from each different point of view. That is because the algorithm is designed to choose the greatest surface with the same descriptor which in some cases like spherical shapes in general, large planes, or other uniform surfaces would find absolute minima in the same shapes that would not reconstruct the object but align incorrect surfaces.

Because of the point simplification process the algorithm is supposed to work in cases where the contour is enough to find correspondences. If the details mark the difference to find correct correspondences, then this approach would not be able to find it correctly, however the assumption of a low definition scan means that the level of detail is not good enough for those cases.

The results of the experiments conducted to the different types of range images showed that as many other previous approaches this solution still finds correspondences for large surfaces of similar curvature even when they are not correct correspondences. This problem can be reduced by the type of scans that are delivered to the program. With a greater area of overlap between the range images, the greatest patch that can be found would be the corresponding overlapping areas of each point cloud. However, the unordered batch alignment for very uniform surfaces like chairs, cars, etc., translates into a greater problem, since possible matches could be found in almost every pair. In these cases the possible solution is to solve the registration in the delivered ordered cycle.

7. FUTURE WORK

The results obtained from this work leave the door open for improvement by using some of the related works as reference. The use of the same structure configuration to parallelize the multi-view range images registration and speed up the process used in [2] as well as the update of the complete structured registered could be adapted to the proposed solution. Another area of opportunity is present in the evaluation metric of the registration; the algorithm uses the ICP or ICRP metrics, however these methods are known to be faulty in the sense that there may exist absolute minimum values that are not the correct solution. The research of an extra constraint to neglect the non-coincident absolute minimum value is a matter of great importance for this and any other work related to the registration.

Lastly, ICP uses a randomized sampling process to select the points that will be optimized for the fine registration, however choosing a small sample the algorithm may receive points outside the overlapping areas of the point clouds, and therefore it becomes necessary to use a greater sample. The use of a large sample makes a very expensive process, but considering that our coarse registration is close enough and that the hierarchical structure is already detecting the intersections between the point clouds, then it could be possible to pass that information so that the sampling is only to be done in the intersecting areas.

ACKNOWLEDGMENTS

We would like to thank Shane Transue who developed the Scannix program for 3D scan, rendering and the ICP implementation.

REFERENCES

- [1] H. Fukai and G. Xu, Fast and Robust Registration of Multiple 3D Point Clouds, *RO-MAN, 2011 IEEE*, pages 331-336, 31 Jul.-3 Aug. 2011.
- [2] A. Makadia; A. Patterson and K. Daniilidis, Fully Automatic Registration of 3D Point

Clouds, *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, vol. 1, pages 1297-1304, 17-22 Jun. 2006.

- [3] J. Wingbermühle, Robust Registration of Coarse Binary Volume Models, *Pattern Recognition, 2000. Proceedings. 15th International Conference on*, vol. 1, pages 999-1002, 3-7 Sep. 2000.
- [4] A.S. Mian; M. Bennamoun and R.A. Owens, "Automatic Multiview Coarse Registration of Range Images for 3D Modeling", *Cybernetics and Intelligent Systems, 2004 IEEE Conference on*, vol. 1, pages 158-163, 1-3 Dec. 2004.
- [5] Jiajing Dai; Jie Yang, A Novel Two-Stage Algorithm for Accurate Registration of 3-D Point Clouds, *Multimedia Technology (ICMT), 2011 International Conference*, pages 6187-6191, 26-28 Jul. 2011.
- [6] Y. Liu, et al, Range Image Registration Using Hierarchical Segmentation and Clustering, *Computational Intelligence in Robotics and Automation (CIRA), 2009 IEEE International Symposium*, pages 328 – 333, 15-18 Dec. 2009.
- [7] P. Krsek, et al, Range Image Registration Driven by a Hierarchy of Surface Differential Features, *22nd Workshop of the Austrian Association for Pattern Recognition 1998*.
- [8] R. Rusu, et al, Fast Point Feature Histograms (FPFH) for 3D Registration, *Robotics and Automation, 2009. ICRA '09. IEEE International Conference on*, pages 3212 - 3217, 12-17 May 2009

ABOUT THE AUTHORS

1. Alejandro Alonso is currently in MS program in Computer Science Department at University of Colorado Denver.
2. Min-Hyung Choi is an associate professor in Computer Science Department at University of Colorado Denver.

BI-ARC SPIRALS IN MINKOWSKI PLANES

Gunter WEISS¹ and Vera WINITZKY de SPINADEL²

¹University of Technology, Vienna, Austria ²University of Buenos Aires, Argentina

ABSTRACT: Starting point of this investigation is the well-known “Golden Spiral”, a GC^1 -bi-arc curve consisting of (Euclidean) quarter circles, which are inscribed into the gnomon squares of a nested set of Golden Rectangles. This curve is self-similar and it is generated by a discrete spiral group Σ applied to one of the quarter circles. Thereby the centres of the circular arcs form an orbit polygon of this group and so do the transition points from one quarter circle to the next. Even one usually visualizes and interprets the spiral group as the product of rotations and dilatations, the construction of the bi-arc spiral can be based on dilatations and translations alone, thus omitting rotations. This idea allows to construct analogues and further generalizations of the nested set of golden rectangles and the Golden Spiral to bi-arc “spirals” in so-called Minkowski planes. Such planes are affine planes endowed with a metric based on a centrally symmetric convex gauge curve, which acts as unit circle. Obviously similar generalisations yield also for Euclidean-circular bi-spirals to general Metallic Means.

Keywords: Minkowski plane, normed plane, circular bi-arc spiral, Golden Mean, Golden Spiral, Metallic Mean.

1. ITRODUCTION 1: GOLDEN SPIRALS

For visualizing the Golden Mean ϕ (and related Mean Values) the standard place of action is the Euclidean plane. In this plane the most common visualization starts with a rectangle the sides of which have length ratio $1:\phi$, ($\phi = 1,618 \dots$). Then, by cutting off a square, one gets a remaining rectangle of the same side length ratio. This new Golden Rectangle is therefore also the result of a similarity transformation $\sigma = \varrho \cdot \delta$, whereby the rotational part ϱ is a quarter turn and the dilatation part δ has factor ϕ^{-1} . Repeating the procedure of cutting off squares and, following the (not mandatory) rule that each new similarity is a power of σ , one receives a nested set of squares inscribed into a nested set of Golden Rectangles, see Figure 1 and e.g. [1], [6], [7], [8], [9], [10]. These sets of squares and rectangles result from the similarity group Σ which is generated by σ . Replacing the squares’ diagonals by quarter circles finally gives the classical GC^1 -smooth “Golden Spiral”.

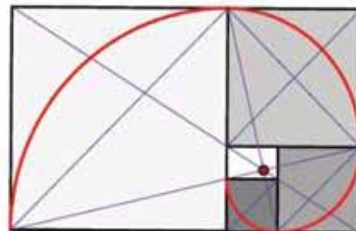


Figure 1: Euclidean Golden Spiral inscribed to a nested set of Golden Rectangles.

With the same justification one could start with a “Golden Triangle” having side length ratio $1:\phi:\phi$ and cut off a gnomon triangle with side length ratio $1:1:\phi$ instead of using a Golden Rectangle and a gnomon square, see Figure 2. Again we can construct a GC^1 -smooth “Golden Spiral” with 108° -circular arc. (Starting with a Golden Triangle of type $1:1:\phi$ and cutting off gnomon triangles of type $1:\phi:\phi$ would deliver another Golden Spiral with 72° -circular arcs.)

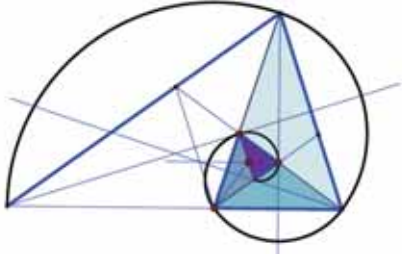


Figure 2: “Golden 108°-circular bi-arc spiral to a nested set of Golden Triangles

Remark 1: When looking at Figures 1 and 2 one might notice that the circular arcs of the spiral also can be understood as translates of certain parts of circles, which are concentric with the start circle. To define such a spiral consisting of circular arcs one just needs a place of action allowing translations and dilatations and with a simply closed, but not necessarily convex curve acting as circles. Therefore the ‘right’ place of action is the *real affine plane*. If we want to talk of circles this affine plane has to be endowed with a norm, thus becoming a Minkowski plane.

2. INTRODUCTION 2: NORMED PLANES

According to the remark above it is sufficient to demand a metric plane as place of action, i.e. an affine plane endowed with a norm, which is induced by a centrally symmetric convex gauge curve as unit circle c . Such a plane is called “Minkowski plane”, see e.g. [11] and [3], [4]. In general the Minkowski norm is defined by a unit circle c , which is neither strictly convex nor smooth. E.g. the well-known “Manhattan metric” as well as the so-called “maximum norm” have (Euclidean) squares, but differently positioned, as unit circles, see [11] and Figure 3. For two points $X_1 = (x_1, y_1)$, $X_2 = (x_2, y_2)$ their “Manhattan distance” is

$d(X_1, X_2) := |(x_1 - x_2) + (y_1 - y_2)|$,
while their maximum norm is

$d(X_1, X_2) := \max(|(x_1 - x_2)|, |(y_1 - y_2)|)$.
For differential geometric purposes one commonly chooses an oval as unit circle. Minkowski

planes comprise also the (abstract) Euclidean plane with an ellipse as unit circle. In general it is not possible to define an orthogonality, which is a symmetric relation between directions. The Euclidean plane and the so-called “Radon planes” are the only ones which have a symmetric orthogonality relation.

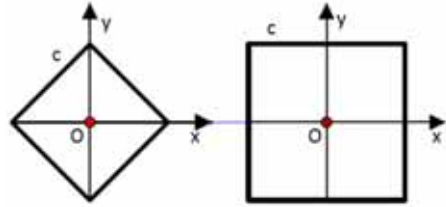


Figure 3: Unit circles c of metric planes:
left c to Manhattan norm,
right c to maximum norm.

Each Minkowski plane allows the full group T of translations τ . As the unit circle c is assumed to be centrally symmetric, also point reflections σ and thus half turns are possible congruence transformations, but in general there are no line reflections and no other rotations than half turns. The group $\Gamma := \text{Aut}(c)$ of affine automorphisms γ of the unit circle c might consist of more than just the identity and half turns. Then $T \circ \Gamma$ will form the group of all congruence transformations of the Minkowski plane. Dilatations δ are, in general, the only similarity transformations in a non-Euclidean Minkowski plane. All similarities of a Minkowski plane can be described as the product $\delta \circ \tau \circ \gamma$ of a dilatation δ , a translation τ and an affine automorphism γ of c .

According to *remark 1* Minkowski planes are possible places of action for visualizing Golden Spirals and their generalizations and in the next section we shall show some special examples of such spirals.

3. EXAMPLES OF GOLDEN SPIRALS IN MINKOWSKI PLANES

We start with a Minkowski unit circle c consisting of four quarter ellipses, see Figure 4. Thereby Euclidean congruence of the four arcs of c as is used in Figure 4 is not essential for constructing a Minkowski spiral. The first step now is to split c arbitrarily into a set of partial arcs. In Figure 5 we used the four given quarter-ellipses as these partial arcs. To these arcs we apply dilatations $\delta^k(\lambda), c_i, k \in \mathbb{Z}$, with a chosen dilatation factor λ and consecutively increasing the power k . As a result we receive a set of arcs c_i , each having start and endpoints S_i, E_i , see Figure 5.

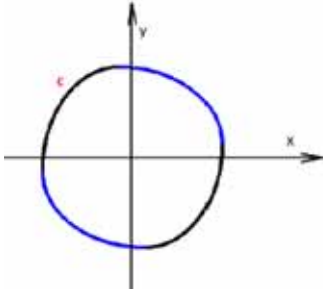


Figure 4: Minkowski unit circle c consisting of congruent quarter circles.

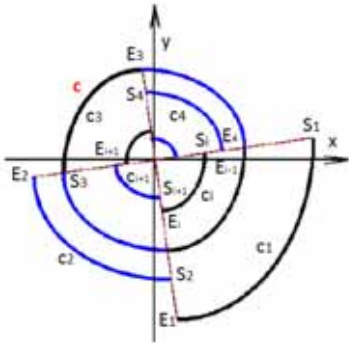


Figure 5: Applying a group of dilatations to a set of Minkowski unit circle arcs.

Finally we apply translations τ_i to these arcs c_i , whereby the translation vector of consecutive arcs is defined by the condition $S_i^{\tau_i} = E_{i-1}^{\tau_{i-1}}$,

see Figure 6. The resulting arcs $c_i^{\tau_i} = d_i$ then are connected at points T_i . We call this condition simply the “end-point is start-point condition”.

Remark: In our example we have chosen the dilatation factor λ such that consecutive arcs d_i, d_{i+1} have GC^3 -transition at T_i . This means that the curvature radius $r_{S_{i+1}}$ of c_{i+1} at vertex S_{i+1} is the same as the curvature radius r_{E_i} of c_i at vertex E_i . If a, b are the semi axes of c_i and $\lambda a, \lambda b$ those of c_{i+1} , then this condition writes as

$$\lambda = \frac{b^3}{a^3}.$$

The additional “closure condition” $1 = \lambda^3 + \lambda^2$ would give a ratio $b:a$ related to van der Laan’s Plastic Number $\psi = 1.32 \dots$, see [2], [5], [12], [13].

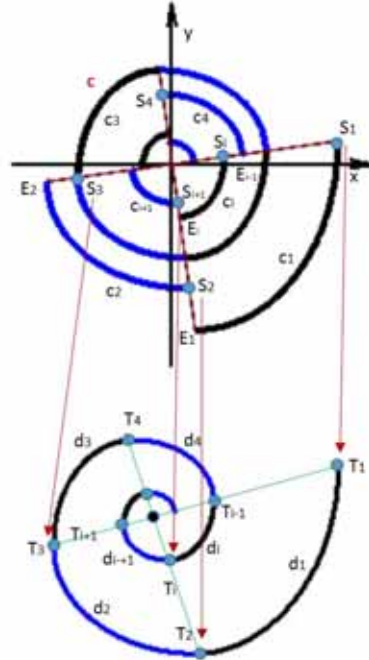


Figure 6: Applying translations τ_i to the arcs c_i under the “end-point is start-point condition” results in a Minkowski circular bi-arc spiral.

When choosing a Euclidean regular k -gon as Minkowski unit circle c , even then the resulting spiral polygon can be called "Minkowski-circular bi-arc spiral", see Fig. 7. In this example we started with an affine regular $2n$ -gon. We splitted it in arcs consisting of two consecutive sides, which at first were dilated and then translated to their final position at a polygonal spiral.

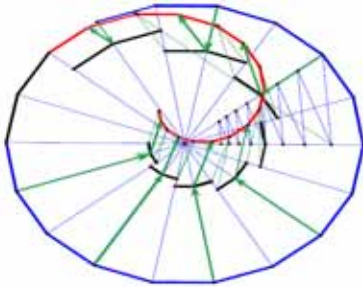


Fig. 7: Minkowski circular bi-arc spiral based on a polygonal Minkowski unit circle.

These two examples show that many of the discrete Euclidean logarithmic spirals treated in references concerning e.g. metallic means spirals and other generalisations (c.f. [7], [8], [9] [10]) have uniform Minkowski geometric interpretations as Minkowski circular bi-arc spirals.

4. FURTHER GENERALISATIONS

The described idea is also applicable to other types of "spirals". E.g. for (discrete) *Archimedean bi-arc spirals* consisting of Minkowski circular arcs one just has to use the set of dilatations $k \cdot \delta(\lambda), k \in \mathbb{Z}$ to an arbitrarily given One can apply the same construction principle and finds relatives to the Euclidean cases in any Minkowski plane, see Figure 8.

In the chapters above we started with a convex, centrally symmetric simply closed curve c as Minkowski unit circle. But, as mentioned in 1. *Remark*, for the construction of the Minkowski spiral convexity of c is not at all essential. One might start with any curved arc c in an affine plane and split it up in infinitely many partial arcs c_i . (Thereby c , if closed, can run through

several or infinitely many times.)

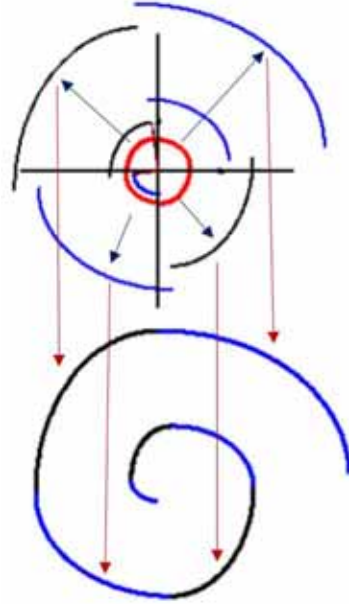


Figure 8: Archimedean Minkowski spiral to Minkowski unit circle c (red) of Figure 4.

Then one chooses a centre O and applies dilatations $\delta^k(\lambda), c_i, k \in \mathbb{Z}$ and $\lambda \in \mathbb{R} \setminus \{0,1\}$ (in the case of log-type spirals) or $k \cdot \delta(\lambda), k \in \mathbb{Z}$ and $\lambda \in \mathbb{R} \setminus \{0\}$ (in the case of Archimedean-type spirals) to these partial arcs c_i . Translations τ_i applied to $c_i^{\delta_i}$ and defined by the "start-point is end-point condition" delivers sort of a "spiral", too. This curve allows, in general, no self-similarities.

In Figure 9 we show an example with a star-shaped base curve c ; its partial arcs c_i are circular arcs between the adjacent vertices of c .

Final Remark: As already stated in [13] it does not really make sense to "model" snail houses, which are objects in space, by (discrete or continuous) log-spirals, which are unlimited and planar mathematical objects. But in spite of this and because of visualisation is limited, too, one might be entrapped to "see" the outlines of real

snail house types in figures like the spiral in Figure 8. The freely chosen base curve c and its splitting in partial arcs c_i together with arbitrarily chosen dilatation factors λ_i gives much freedom for such a rough approximation of real objects.

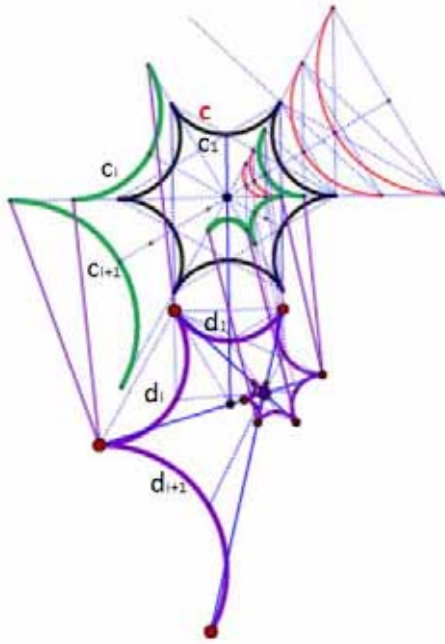


Figure 8: Spiral to a non-convex base curve c .

5. CONCLUSIONS

We have shown that the right place of action for classical visualizations of the Golden Spiral and its generalizations is the real affine plane, as it is possible to construct such spirals by dilatations and translations alone. Applying dilatations with factors λ_i according to a geometric progression to the partial curves c_i one will get a generalized (discrete) affine log-spiral. If the λ_i follow an arithmetic sequel, the resulting spiral will be a generalized (discrete) affine Archimedian Spiral. In an affine plane endowed with a norm (= Minkowski plane), where the norm is given by a convex and centrally symmetric gauge curve c

as “unit circle”, there exists the concept “circle”. This allows to call a wide range of discrete log- and Archimedian spirals “circular bi-arc spirals in the sense of Minkowski”.

REFERENCES

- [1] A. Beutelspacher, B. Petri: *Der Goldene Schnitt*. 2nd ed., Spektrum-Akademischer Verlag 1997
- [2] H. v.d. Laan: *Le Nombre Plastic; quinze leçons sur l'ordonnance archi-tectonique*. Brill, Leiden 1960.
- [3] H. Martini, K. Swanepoel, G. Weiss: *The geometry of Minkowski spaces - a survey*, Part I, *Expositiones Mathematicae* 19 (2001) 97-142, ISSN 0723-0869
- [4] H. Martini, M. Spirova: *Golden rectangles in normed planes*. *Mitteilungen. d. Math. Ges. Hamburg* 29(2010), p.125-134 (MR 2012 a: 51019)
- [5] R. Padovan: *Dom Hans van der Laan and the Plastic Number*. *Nexus Network Journal* Vol.IV(3), *Architecture and Mathematics* (2002), p.181-193 [pdf].
- [6] V.W. de Spinadel: *Golden and Metallic Means in Modern Mathematics and Physics*. *Proc. 13th ICGG 2008*, Dresden, Germany, ISBN 978-3-86780-042-6.
- [7] V.W. de Spinadel: *Metallic Spirals*. *Proc. 10th Int. Conf. APLIMAT 2011*, Bratislava, Slovakia, p.659-666.
- [8] V.W.de Spinadel, L.L.Echagüe: *Espirales de Plata*. *Journ. Math.& Design*, Vol.8 (2008). ISSN 1515.7881
- [9] V.W. de Spinadel, A. Redondo Buitrago: *Towards van der Laan's Plastic Number in the plane*. *JGG* Vol. 13(2), (2008), p.163-175.
- [10] V.W. de Spinadel, A. Redondo Buitrago: *Generalizing the Golden Spiral*. *Journ. Math. & Design* Vol. 11(1)(2011), ISSN 1515.7881.

- [11] A.C.Thompson: *Minkowski Geometry*. Cambridge Univ. Press 1996, ISBN 0-521-40472-X
- [12] G. Weiss, V.W. de Spinadel: *Remarks to classical cubic problems and the mean values of van der Laan and Rosenbusch*. Proc. 14th ICGG 2010, Kyoto, Japan, Art. 235, p. 1-11. ISBN 987-4-9900967-1-7
- [13] G.Weiss: From the Golden Rectangle to the Laan Box and more. Journ. Math. & Design Vol. 12 (2012), ISSN 1515.7881.

ABOUT THE AUTHORS

1. Gunter WEISS, Dr. is a retired professor for Geometry (since 2011) at the Dresden University of Technology, Germany. His e-mail address is weissgunter@hotmail.com .
2. Vera Winitzky de Spinadel, Dr., is a Full Emeritus Professor (since 2010) of mathematics at the University of Buenos Aires; director of the research Centre of Mathematics & Design (MAyDI); president of the International Mathematics & Design Association since 1998; chief editor of the Journal of Mathematics & Design.
Her e-mail is vspinade@fibertel.com.ar and myd_lab@yahoo.com.ar .

BRAUNER'S ANGLE FORMULA AND THE THEORY OF CURVES IN MINKOWSKI PLANES AND SPACES

Emad N. SHONODA^{1,2}, and Gunter WEISS³

¹Aljouf University, Saudi Arabia, ²Port-Said University, Egypt

³University of Technology, Vienna, Austria

ABSTRACT: A Minkowski space is a normed (affine) space, whereby the metric is based on a convex and centrally symmetric unit ball \mathbf{B} . If this ball is not an ellipsoid, such a space is not an inner product space, such that angle measure and (geometrically reasonable) orthogonality concepts are not symmetric relations. A very commonly used orthogonality concept is due to Birkhoff (see [4]) and we will also use this concept throughout the paper distinguishing between “left” and “right orthogonality”. Following an idea of H. Brauner [1] it is possible to define an angle measure via the directions of the angle’s legs and their orthogonal directions (in the plane of that angle). We apply this concept to Minkowski planes and spaces and can show connections to so-called Minkowski angle functions like the “cosine-Minkowski”. Following H. Guggenheimer [2] it is possible to define a left-orthogonal moving frame of (at first a planar) differentiable curve based on B-orthogonality and translate classical differential geometric concepts of curves into Minkowski Geometry. Thereby it is possible to formulate an analogue to the “main theorem” of planar curves. We also present a Minkowski geometric version of Serret-Frenet formulae in a Minkowski plane. Furthermore we can adjust a result of H. Guggenheimer [2] based on Minkowski trigonometric functions. Finally also space curves are treated; here one has to distinguish four curvatures and torsions.

Keywords: Minkowski spaces, Brauner’s angle formula, Minkowski differential geometry of curves.

1. INTRODUCTION

Normed spaces are real affine spaces endowed with a metric based on a centrally symmetric unit ball \mathbf{B} . Such “MINKOWSKI-spaces” are, in general, not inner product spaces and therefore one has to modify the classical concepts of orthogonality and angle measure. There are many attempts to define “Minkowski-orthogonality”, most of them justified by and applied to just a single geometric problem and only exceptionally they are symmetric relations. Most commonly used is “BIRKHOFF-orthogonality (B-orthogonality) with the disadvantage of being a non-symmetric relation. One has to distinct “left-orthogonal” \dashv and

“right-orthogonal” \vdash directions, see Figure 1 for the planar case.

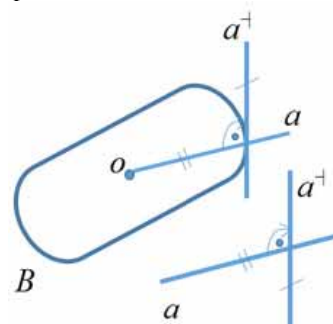


Figure 1: BIRKHOFF-left-orthogonality in a Minkowski plane with unit ball \mathbf{B} .

The most suitable of the many orthogonality definitions in a Minkowski space M_B^n is due to Birkhoff which state that, if M_B^n is an n -dimensional normal linear space "Minkowski space" of unit ball \mathbf{B} and $\mathbf{x}, \mathbf{y} \in M_B^n$ then we say that $\mathbf{x}, \mathbf{y} \in M_B^n$

$$\mathbf{x} \perp \mathbf{y} \Leftrightarrow \|\mathbf{x}\| \leq \|\mathbf{x} + \lambda \mathbf{y}\| \quad \forall \lambda \in \mathbb{R} \quad (1)$$

For differential geometric purposes it is necessary that Birkhoff-orthogonality is a one-to-one relation. Therefore the following investigations are based on the assumption: The centric symmetric gauge ball \mathbf{B} of a Minkowski plane or space is assumed to be strictly convex and (at least) C^1 -smooth.

Birkhoff "B-orthogonality" is, in general, a non symmetric relation. However, it is symmetric in the cases of a Euclidean or Radon plane, c.f. [4]. In those two cases all B-orthogonal vectors are mutually normal. In the general (planar) case there is at least one pair of directions¹, where B-orthogonal directions are mutually orthogonal. We will use a coordinate frame $\{O, \mathbf{x}, \mathbf{y}\}$ in a Minkowski plane based on such a pair of mutually B-orthogonal directions. This allows to formulate some results concerning the so-called *semi-inner product* of two vectors in M-plane, see [7].

Further references for Minkowski geometry are e.g. [20], [1], [2], [11], [12], [14-16].

2. BRAUNER'S ANGLE FORMULA

A famous formula of E. Laguerre (1853) describes the (Euclidean) angle between two lines a, b by means of Projective Geometry as follows: Let Π the projectively enclosed Euclidean plane and let A, B be the (then existing)

ideal points of a, b and I, J the pair of conjugate imaginary "absolute points" on the ideal line $u = A \vee B$. These points I, J together with line u form the (singular quadratic) "absolute figure" of planar Euclidean geometry. According to Laguerre the angle measure $\varphi := \angle \mathbf{a} \mathbf{b}$ calculates as

$$\varphi = \angle \mathbf{a} \mathbf{b} = \frac{i}{2} \ln (cr(IJAB)). \quad (2)$$

This formula has, in spite of its importance for understanding Euclidean Geometry as a sub-geometry of Projective Geometry, two disadvantages: The first is the necessity of a complex extension of the places of action with the consequence that one has to deal with non ordered angles. The second is the necessity of a quadratic absolute figure, thus demanding that the underlying vector space has to be an inner product space.

H. BRAUNER's angle formula (1976) uses instead of Laguerre's "absolute points" I, J the ideal points $A', B' \in u$ belonging to the direction vectors \mathbf{a}', \mathbf{b}' orthogonal to the directions \mathbf{a}, \mathbf{b} . It reads as follows:

$$\begin{aligned} \tan(\varphi) &:= \sqrt{-cr(A', B, B')} = \\ &= \sqrt{-cr(\mathbf{a}, \mathbf{a}', \mathbf{b}, \mathbf{b}')}, \end{aligned} \quad (3)$$

with the addendum that, if \mathbf{a} orthogonal \mathbf{b} (meaning $\mathbf{a}' = \mathbf{b} \wedge \mathbf{b}' = \mathbf{a}$ such that in the essential function "cross-ratio" only occur two different elements instead of four) then $\tan \varphi = \frac{\pi}{2}$. This formula avoids complex extension and it is based on an orthogonality structure of the place of action. It is therefore possible to declare an angle concept with this formula in a large set of normed spaces and connect it to concepts already in use, like "cosine-Minkowski". In fact one has many possibilities to do this according to the many different orthogonality structures known for a Minkowski plane.

¹In the following we use the term "direction" and its symbols for lines as well as for the pencil of parallel lines. As we deal with the projective extension of Euclidean and Minkowski spaces, orientation of lines has no meaning, but comes in via the analytic description with respect to affine coordinate systems. This justifies to use the concept "direction vector" as well.

Given two directions \mathbf{a}, \mathbf{b} one constructs the left-orthogonal directions $\mathbf{a}^\perp, \mathbf{b}^\perp$, see Figure 2, then we define “tangents-Minkowski” tm as

$$tm \angle_M \mathbf{a} \mathbf{b} := \sqrt{-cr(\mathbf{a}, \mathbf{a}^\perp, \mathbf{b}, \mathbf{b}^\perp)}. \quad (4)$$

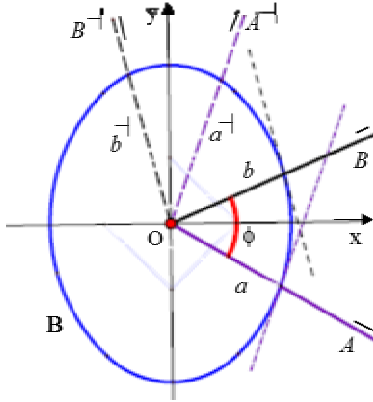


Figure 2: Adaption of H. Brauner's angle formula to a Minkowski plane based on Birkhoff's left-orthogonality to a strictly convex gauge ball \mathbf{B} .

One has to discuss possible cases of mutual positions of $\mathbf{a}, \mathbf{b}, \mathbf{a}^\perp, \mathbf{b}^\perp$, to give a complete “Minkowski-Brauner definition” of an angle measure in a Minkowski plane, as now there are cases with only three different directions and, as for Euclidean orthogonal lines, even only two.

Remark: The essential function in (4) is the cross ratio and the properties of tm follow from both, the properties of cross ratio and the orthogonality structure, which, in our case is Birkhoff orthogonality. From the permutation law of cross ratio (e.g. [9, p.36] follows e.g. that

$$\angle_M \mathbf{a} \mathbf{b} = \angle_M \mathbf{b} \mathbf{a}, \quad (5)$$

while B-orthogonality (in general) is responsible for the angle measure not being additive.

3. M-TRIGONOMETRIC FUNCTIONS IN MINKOWSKI-SPACES

A very general and well suited definition of the *cosine Minkowski function* cm between two directions \mathbf{x}, \mathbf{y} in M_B^n can be found e.g. in [13].

The two directions vectors \mathbf{x}, \mathbf{y} span a diameter plane of \mathbf{B} , which intersects \mathbf{B} in a centric symmetric, smooth and strictly convex curve \mathbf{B}_2 , which makes this plane to a Minkowski plane $M_{B_2}^2$.

We assume $\mathbf{B} = \{\mathbf{x} / \|\mathbf{x}\|_M \mid \mathbf{x} \in M_B^n\}$ to be smooth and strictly convex. Then there exists a linear function f_x , which is unique up to a (positive) scalar factor and which for all $\mathbf{x} \in M_B^n, \mathbf{x} \neq 0$ attains its maximum at \mathbf{x} :

$$f_x(\mathbf{x}) = \|\mathbf{x}\| \cdot \|f_x\| \quad (6)$$

Using this function we can define the M-cosine function as follows:

Definition 3.1. For an ordered pair of tors $\mathbf{x}, \mathbf{y} \in M_B^n, \mathbf{y} \neq 0$, the cosine- Minkowski function is the (non symmetric) function $cm(\mathbf{x}, \mathbf{y})$ with

$$cm(\mathbf{x}, \mathbf{y}) := \frac{f_x(\mathbf{y})}{\|\mathbf{y}\| \cdot \|f_x\|}. \quad (7)$$

Substituting (6) into (7) we get

$$cm(\mathbf{x}, \mathbf{y}) = \frac{\|\mathbf{x}\| \cdot f_x(\mathbf{y})}{\|\mathbf{y}\| \cdot f_x(\mathbf{x})} \quad (8)$$

and we can extend the definition for a pair (vector \mathbf{x} , subspace $M^d \subset M_B^n$) as follows:

$$cm(\mathbf{x}, M^d) := \max \{cm(\mathbf{x}, \mathbf{y}) \mid \mathbf{y} \in M^d \setminus \{0\}\}$$

Also a *sine-Minkowski* function can be defined: For this we need the dual space $(M_B^n)^*$ of a

Minkowski space M_B^n and a norm of this space, too. Such a norm function $\tilde{\sigma}$ is induced by the so-called isoperimetric \tilde{I}_B of B in $(M_B^n)^*$, see e.g. [20].

Definition 3.2. Let $H = \text{span}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{n-1})$ be a hyperplane through the origin of M_B^n and let $\mathbf{x} \in M_B^n$ be a non-zero vector; suppose that there exists a linear function f in the dual space $(M_B^n)^*$ of M_B^n such that $f^\perp = H$ and $\text{sgn}(f) := \text{sgn}(\det(\mathbf{x}_1, \dots, \mathbf{x}_{n-1}, \mathbf{x}))$, then the *sine-Minkowski* function sm is declared as

$$sm(H, \mathbf{x}) := \frac{f(\mathbf{x})}{\|\mathbf{x}\| \cdot \tilde{\sigma}(f)}. \quad (10)$$

Many interesting properties of cm and sm connected with B-orthogonality can be found in [20] and [13].

Definition 3.3. Let $\mathbf{x}_1, \mathbf{x}_2$ be two vectors in a Minkowski plane M_B^2 , then this pair is called *mutually normal* pair, if $\mathbf{x}_1 \perp \mathbf{x}_2$ and $\mathbf{x}_2 \perp \mathbf{x}_1$. Normed mutually normal vectors are called *orthonormal*.

Proposition 3.1. ([20]): Let M_B^n be a Minkowski space with unit ball B . Then there exists at least one basis $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$ such that $\|\mathbf{x}_i\| = 1$ and $\mathbf{x}_i \perp \mathbf{x}_j$ for all i and j with $i \neq j$; i.e. each pair of basis vectors is an orthonormal pair.

Proposition 3.2. ([20]): If one can construct an orthonormal basis to any start vector $\mathbf{x} \in M_B^n$, $\mathbf{x} \neq 0$, then M_B^n is Euclidean for $n \geq 3$ and Euclidean or a Radon plane for $n = 2$.

Definition 3.4. (*B-orthonormal frame* in M_B^n): Let $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n \in M_B^n$, $\|\mathbf{e}_i\| = 1 \forall i = 1, 2, \dots, n$. If $\mathbf{e}_j \perp \mathbf{e}_k \forall k = 1, 2, \dots, j-1$ and $j = 2, 3, \dots, n$, then the ordered vector set $(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$ is

called *B-orthonormal frame* in M_B^n .

4. SEMI-INNERPRODUCT IN MINKOWSKI SPACES

A normed linear space is not necessarily an inner product space. However, a real normed linear space is an inner product space if and only if each two-dimensional linear subspace of it is an inner product space. Equivalently, we can say that a normed linear space is an inner product space if and only if every plane section of the unit ball B passing through the origin O is an ellipse. For $n \geq 3$, B is an ellipsoid and the Minkowski space is Euclidean. As we cannot start with an inner product, we need to find a replacement, which is compatible with B-orthogonality concept and the non-Euclidean Minkowski norm. It is called "*semi-inner product*", see e.g. [7], [8], [20].

Theorem 4.1. ([8]): In each Minkowski space M_B^n there exists at least one well defined semi-inner product $\langle \dots \rangle_M$ which generates the norm $\|\cdot\|$. That is, $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x} \cdot \mathbf{x} \rangle}$ for all $\mathbf{x} \in M_B^n$, and it is unique if and only if M_B^n is smooth and strictly convex. Based on f_x (6) this semi-inner product is declared by

$$\langle \mathbf{x}, \mathbf{y} \rangle_M := \frac{f_x(\mathbf{y})}{f_x(\mathbf{x})} \|\mathbf{x}\|^2. \quad (11)$$

Substituting (7) into (11) we finally get

$$\langle \mathbf{x}, \mathbf{y} \rangle_M = \|\mathbf{x}\| \|\mathbf{y}\| cm(\mathbf{x}, \mathbf{y}). \quad (12)$$

For properties of the semi-inner product (11) and its relation to B-orthogonality we refer to [16] and [18].

5. BRAUNER'S ANGLE MEASURE AND M-TRIGONOMETRIC FUNCTIONS

Based on our modification of Brauner's angle measure (4) and the definition of cm we extend the Minkowski plane to its projective enclosure and can state the following

Theorem 5.1. ([18]): If $\mathbf{u}, \mathbf{v} \in M_B^2$ are two non-zero vectors, then the angle between \mathbf{u} and \mathbf{v} satisfies the equation

$$cm(\mathbf{u}, \mathbf{v}) \cdot cm(\mathbf{v}, \mathbf{u}) = cr(\mathbf{u}, \mathbf{v}, \mathbf{v}^\perp, \mathbf{u}^\perp). \quad (13)$$

We can rewrite (13) using P. Milićić's notation (see [16]) as follows:

$$\overline{cm}(\mathbf{u}, \mathbf{v}) := \sqrt{cr(\mathbf{u}, \mathbf{v}, \mathbf{v}^\perp, \mathbf{u}^\perp)} \quad (14)$$

Equations (13) and (14) give a relation between a pair of oriented angles and the cross ratio of its vectors and their B-orthogonal ones. In the following theorem we proof that the function \overline{cm} (14) has very nice properties.

Theorem 5.2. Given an orthonormal frame $\{O, \mathbf{a}, \mathbf{b}\}$ of M_B^2 , then for any non-zero vector $\mathbf{x} \in M_B^2$ and the coordinate vectors \mathbf{a}, \mathbf{b} yields

$$(\overline{cm}(\mathbf{a}, \mathbf{x}))^2 + (\overline{cm}(\mathbf{b}, \mathbf{x}))^2 = 1. \quad (15)$$

This formula can be proved in two ways. First by using Brauner's formula in Minkowski space, second, we can prove it by using M- trigonometric functions. This means the Brauner's theorem in M-space is well defined.

Proof 1. We begin with Brauner's formula (3) and (13). (Thereby we formally use the ideal points A_∞, \dots of the lines $\mathbf{a}, \mathbf{b}, \mathbf{x}$, and their B-orthogonal directions.) Substituting (14) into the left hand side of (15) delivers:

$$\overline{cm}(\mathbf{a}, \mathbf{x})^2 + \overline{cm}(\mathbf{b}, \mathbf{x})^2 = cr(A_\infty, X_\infty, X_\infty^\perp, A_\infty^\perp) + cr(B_\infty, X_\infty, X_\infty^\perp, B_\infty^\perp).$$

Since (\mathbf{a}, \mathbf{b}) is a mutually normal basis of M_B^2 , i.e. $A_\infty^\perp = B_\infty$ and $B_\infty^\perp = A_\infty$, we have

$$\begin{aligned} cr(A_\infty, X_\infty, X_\infty^\perp, A_\infty^\perp) &= cr(A_\infty, X_\infty, X_\infty^\perp, B_\infty) \\ &= 1 - cr(A_\infty, X_\infty^\perp, X_\infty, B_\infty) \\ &= 1 - cr(B_\infty, X_\infty, X_\infty^\perp, A_\infty). \end{aligned} \quad (16)$$

such that we indeed get (15). ■

Proof 2. Now we try to prove theorem 5.2 by using M- trigonometric functions: We apply the following statements, see [20]:

Corollary 5.1. ([20; p.263]): For any non-zero vectors $\mathbf{x}_1, \mathbf{x}_2 \in M_B^2$ we have

$$\begin{aligned} (\overline{cm}(\mathbf{x}_1, \mathbf{x}_2))^2 + \frac{sm(\mathbf{x}_1, \mathbf{x}_2) \cdot sm(\mathbf{x}_1^\perp, \mathbf{x}_2^\perp)}{\alpha(\mathbf{x}_1^\perp) \alpha(\mathbf{x}_2^\perp)} \\ = 1 \end{aligned} \quad (17)$$

where the function α is given by

$$\alpha(\mathbf{x}) := \sup\{sm(H, \mathbf{x}) : H \subset M_B^2\} = \|\mathbf{x}\|_I / \|\mathbf{x}\|.$$

If we assume that $\mathbf{x}_1 = \mathbf{a}$ and $\mathbf{x}_2 = \mathbf{x}$, then (17) becomes

$$(\overline{cm}(\mathbf{a}, \mathbf{x}))^2 + \frac{sm(\mathbf{a}, \mathbf{x}) \cdot sm(\mathbf{b}, \mathbf{x}^\perp)}{\alpha(\mathbf{b}) \alpha(\mathbf{x}^\perp)} = 1 \quad (18)$$

Corollary 5.2: ([20; p.237] For any non-zero vectors $\mathbf{x}_1, \mathbf{x}_2 \in M_B^2$ yields

$$sm(\mathbf{x}_1^\perp, \mathbf{x}_2) = \alpha(\mathbf{x}_1^\perp) \cdot cm(\mathbf{x}_1, \mathbf{x}_2) \quad (19)$$

Again, we assume that $\mathbf{x}_1 = \mathbf{b}, \mathbf{x}_2 = \mathbf{x}$, then (19) becomes

$$sm(\mathbf{a}, \mathbf{x}) = \alpha(\mathbf{a}) cm(\mathbf{b}, \mathbf{x}) \quad (20)$$

and putting $\mathbf{x}_1 = \mathbf{x}$ and $\mathbf{x}_2 = \mathbf{b}$ we get

$$sm(\mathbf{x}^\perp, \mathbf{b}) = -\alpha(\mathbf{x}^\perp) \cdot cm(\mathbf{x}, \mathbf{b}) \quad (21)$$

where the sign in (21) comes from the orientation of the angle because of

$$sm(\mathbf{x}^\perp, \mathbf{b}) = -sm(\mathbf{b}, \mathbf{x}^\perp). \quad (22)$$

Substitution of (22) into (20) and of (19) into (17) and taking into account that $\alpha(\mathbf{a}) = \alpha(\mathbf{b}) = 1$, we can rewrite (18) as

$$cm(\mathbf{a}, \mathbf{x}) cm(\mathbf{x}, \mathbf{a}) + cm(\mathbf{b}, \mathbf{x}) cm(\mathbf{x}, \mathbf{b}) = 1. \quad \blacksquare$$

Theorem 5.3. Let $\mathbf{x}, \mathbf{y}, \mathbf{z} \in M_B^2$ be three directions. Their ideal points are $X_\infty = (0, \mathbf{u})\mathbb{R}$, $Y_\infty = (0, \mathbf{v})\mathbb{R}$, $Z_\infty = (0, \mathbf{w})\mathbb{R}$ respectively. Then $\mathfrak{z}_M(\mathbf{u}, \mathbf{v}) = \mathfrak{z}_M(\mathbf{v}, \mathbf{w})$, (which is equivalent to $\overline{cm}(\mathbf{u}, \mathbf{v}) = \overline{cm}(\mathbf{v}, \mathbf{w})$), if and only if the following condition is valid

$$cr(X_\infty, Y_\infty, Y_\infty^\perp, X_\infty^\perp) = cr(Y_\infty, Z_\infty, Z_\infty^\perp, Y_\infty^\perp) \quad (23)$$

Proof. The proof is obvious as the condition simply relates to Brauner's formula (14). It might be of interest to reformulate (23) by using such representations of the direction vectors $\mathbf{u}, \mathbf{v}, \dots$ describing ideal points, $X_\infty, Y_\infty, \dots$ having 1 (or 0) as the first coordinate. Then one can write (23) as

$$\frac{v^\perp - u}{v^\perp - v} : \frac{u^\perp - u}{u^\perp - v} = \frac{w^\perp - v}{w^\perp - w} : \frac{v^\perp - v}{v^\perp - w}$$

from what follows

$$\frac{v^\perp - u}{v^\perp - w} = \frac{u^\perp - u}{u^\perp - v} \cdot \frac{w^\perp - v}{w^\perp - w} \quad (24)$$

where u, v, \dots are the second homogeneous coordinates of the ideal points $X_\infty, Y_\infty, \dots$ respectively. Equation (24) could be interpreted as a condition for "ratios" of triplets of ideal points. ■

From theorem 5.3 follows that if (\mathbf{x}, \mathbf{y}) is an orthonormal basis of M_B^2 and \mathbf{z} fulfils condition (23), then $\mathbf{z} = -\mathbf{x}$. Furthermore follows

Corollary 5.3. If $\mathbf{x}^\perp = \mathbf{y}$ and $\mathbf{z} = -\mathbf{x}$ then (\mathbf{x}, \mathbf{y}) is an orthonormal basis of M_B^2 . If this is true for all such $\mathbf{x}, \mathbf{y} \in M_B^2$ then the plane is Radon or Euclidean plane.

Theorem 5.4. The function $\overline{cm}(\mathbf{u}, \mathbf{v})$, $\forall \mathbf{u}, \mathbf{v} \in M_B^2$, has the following properties:

- i. $\overline{cm}(\mathbf{u}, \mathbf{v}) = 1 \Leftrightarrow \mathfrak{z}_M(\mathbf{u}, \mathbf{v}) = 0$.
- ii. $\overline{cm}(\mathbf{u}, \mathbf{v}) = \overline{cm}(-\mathbf{u}, -\mathbf{v})$
 $\Leftrightarrow \mathfrak{z}_M(\mathbf{u}, \mathbf{v}) = \mathfrak{z}_M(-\mathbf{u}, -\mathbf{v})$.
- iii. cosine- and sine-Minkowski functions are related by

$$\left(\overline{cm}(\mathbf{u}, \mathbf{v})\right)^2 = \|\mathbf{u}^\perp\| \|\mathbf{v}^\perp\| \left(\overline{sm}(\mathbf{u}, \mathbf{v})\right)^2 \quad \forall \mathbf{u}, \mathbf{v} \in M_B^2$$
where $\overline{sm}(\mathbf{u}, \mathbf{v}) = \sqrt{\overline{sm}(\mathbf{u}^\perp, \mathbf{v}) \overline{sm}(\mathbf{v}^\perp, \mathbf{u})}$.

The proofs of these properties are obvious consequences of Theorem 5.3 and Corollary 5.1 and 5.2.

6. PLANE CURVES IN MINKOWSKI PLANE

In the following we apply the above developed concepts of M-trigonometric functions and their connection to Brauner's formula to Elementary Differential Geometric topics in Minkowski planes and later also in Minkowski-3-spaces.

Consider a plane curve $\gamma: [a, b] \rightarrow M_B^2$ which is given by two coordinate functions

$$\gamma(x(t), y(t)), \quad t \in [a, b].$$

The curve γ is assumed to be regular, i.e. $\gamma' = \frac{d\gamma}{dt} \neq 0$ in $[a, b]$.

The unit tangent vectors \mathbf{T} of γ and the normed normal vectors \mathbf{N} constitute Frenet's tangent vector field resp. normal vector field. With respect to B-orthogonality \mathbf{N} has to be defined as right-orthogonal to \mathbf{T} , i.e. $\mathbf{T} \dashv \mathbf{N}$

Moreover, we can construct a vector $\mathbf{N}^\perp \in \widetilde{I}_B$ with $\mathbf{N} \dashv \mathbf{N}^\perp$. Then we have two normal bases,

(\mathbf{T}, \mathbf{N}) and $\left(\mathbf{N}, \frac{\mathbf{N}^\perp}{\|\mathbf{N}^\perp\|}\right)$ in plane M_B^2 . Besides

these bases there exists also (at least) one orthonormal basis (\mathbf{a}, \mathbf{b}) , which will be used as a fixed coordinate frame $\{O, \mathbf{a}, \mathbf{b}\}$. Since, $\mathbf{T} \dashv \mathbf{N}$ and $\mathbf{N} \dashv \mathbf{N}^\perp$, we can get the derivative of the unit vectors \mathbf{T} and \mathbf{N} as follows:

$$\begin{aligned}\mathbf{T}' &= p\mathbf{T} + q\mathbf{N}, \\ \mathbf{N}' &= \alpha\mathbf{N} + \beta \frac{\mathbf{N}^\perp}{\|\mathbf{N}^\perp\|}, \quad p, q, \alpha, \beta \in \mathbb{R}.\end{aligned}\quad (25)$$

Using an “arc length parameter” one can get $p = 0$ and then put $q = : \chi$. Furthermore, \mathbf{N}^\perp must be a linear combination of the normal basis $(\mathbf{T}, \mathbf{T}^\perp)$. Therefore, we can find the second Frenet-Serret formula. We use a representation of the vectors with respect to the fixed orthonormal frame (\mathbf{a}, \mathbf{b}) and the (affine) coordinate transformation from (\mathbf{a}, \mathbf{b}) to $(\mathbf{T}, \mathbf{T}^\perp)$. Thereby we put

$$\mathbf{N}^\perp = \lambda_1 \mathbf{a} + \lambda_2 \mathbf{b}, \quad \lambda_i \in \mathbb{R} \quad (26)$$

Applying semi-inner products we can calculate the factors λ_1, λ_2 as

$$\begin{aligned}\lambda_1 &= \|\mathbf{N}^\perp\| \, cm(\mathbf{a}, \mathbf{N}^\perp) \\ \lambda_2 &= \|\mathbf{N}^\perp\| \, cm(\mathbf{b}, \mathbf{N}^\perp).\end{aligned}$$

With the abbreviation $\hat{\mathbf{N}}^\perp = \mathbf{N}^\perp / \|\mathbf{N}^\perp\|$ follows therefore

$$\hat{\mathbf{N}}^\perp = cm(\mathbf{a}, \mathbf{N}^\perp) \mathbf{a} + cm(\mathbf{b}, \mathbf{N}^\perp) \mathbf{b}. \quad (27)$$

Corollary 6.1. ([20, Prop. 8.3.5]): With respect to two normal bases $(\mathbf{T}, \mathbf{T}^\perp)$ and (\mathbf{a}, \mathbf{b}) the mapping $\tau: M_B^2 \rightarrow M_B^2; (\mathbf{a}, \mathbf{b})' \rightarrow (\mathbf{T}, \mathbf{T}^\perp)'$ is described by the matrix

$$\psi = \begin{pmatrix} cm(\mathbf{T}, \mathbf{a}) & sm(\mathbf{T}, \mathbf{a}) \\ -\xi sm(\mathbf{T}^\perp, \mathbf{b}) & cm(\mathbf{a}, \mathbf{T}) \end{pmatrix}, \quad (28)$$

where $\xi = \alpha(\mathbf{a})^{-1} \alpha(\mathbf{T}^\perp)^{-1}$.

Substituting (28) into (27) gives

$$\begin{aligned}\hat{\mathbf{N}}^\perp &= cm(\mathbf{a}, \mathbf{N}^\perp) (cm(\mathbf{T}, \mathbf{a}) \mathbf{T} + sm(\mathbf{T}, \mathbf{a}) \mathbf{T}^\perp) + \\ &\quad cm(\mathbf{b}, \mathbf{N}^\perp) (-\xi sm(\mathbf{T}^\perp, \mathbf{b}) \mathbf{T} + cm(\mathbf{a}, \mathbf{T}) \mathbf{b})\end{aligned}$$

and simplified this is

$$\begin{aligned}\hat{\mathbf{N}}^\perp &= \{cm(\mathbf{a}, \mathbf{N}^\perp) cm(\mathbf{T}, \mathbf{a}) - \\ &\quad \xi cm(\mathbf{b}, \mathbf{N}^\perp) sm(\mathbf{T}^\perp, \mathbf{b})\} \mathbf{T} + \\ &\quad \{cm(\mathbf{a}, \mathbf{N}^\perp) sm(\mathbf{T}, \mathbf{a}) + \\ &\quad cm(\mathbf{b}, \mathbf{N}^\perp) cm(\mathbf{a}, \mathbf{T})\} \mathbf{b}.\end{aligned}\quad (29)$$

Using Corollary 5.2 in the form $\mathbf{x}_1, \mathbf{x}_2 \in M_B^2$
 $\Rightarrow sm(\mathbf{x}_1, \mathbf{x}_2) = \alpha(\mathbf{x}_1) cm(\mathbf{x}_1, \mathbf{x}_2)$

and assuming that $\mathbf{a}^\perp = \mathbf{b}$ and $\mathbf{b}^\perp = \mathbf{a}$ follows:

$$\begin{aligned}\hat{\mathbf{N}}^\perp &= \{cm(\mathbf{a}, \mathbf{N}^\perp) cm(\mathbf{T}, \mathbf{a}) - \\ &\quad \xi sm(\mathbf{a}, \mathbf{N}^\perp) sm(\mathbf{T}^\perp, \mathbf{b})\} \mathbf{T} + \\ &\quad \{cm(\mathbf{a}, \mathbf{N}^\perp) sm(\mathbf{T}, \mathbf{a}) + \\ &\quad sm(\mathbf{a}, \mathbf{N}^\perp) cm(\mathbf{a}, \mathbf{T})\} \mathbf{T}^\perp.\end{aligned}\quad (30)$$

where, $sm(\mathbf{a}, \mathbf{N}^\perp) = cm(\mathbf{b}, \mathbf{N}^\perp)$ and

$$\alpha(\mathbf{a}) = \alpha(\mathbf{b}) = 1.$$

Applying Corollary 6.1 we finally get

$$\hat{\mathbf{N}}^\perp = cm(\mathbf{T}, \mathbf{N}^\perp) \mathbf{T} + sm(\mathbf{T}, \mathbf{N}^\perp) \mathbf{T}^\perp. \quad (31)$$

Substituting (31) in (27) we get,

$$\mathbf{N}' = -\chi (cm(\mathbf{T}, \mathbf{N}^\perp) \mathbf{N} + sm(\mathbf{T}, \mathbf{N}^\perp) \mathbf{T}^\perp), \quad (32)$$

which is the second Frenet-Serret formula. Therefore, the Frenet matrix could be written as follows

$$\begin{pmatrix} \mathbf{T}' \\ \mathbf{N}' \end{pmatrix} = \begin{pmatrix} 0 & \chi \\ -\chi cm(\mathbf{T}, \mathbf{N}^\perp) & -\chi sm(\mathbf{T}, \mathbf{N}^\perp) \end{pmatrix} \begin{pmatrix} \mathbf{T} \\ \mathbf{T}^\perp \end{pmatrix} \quad (33)$$

whereby $\mathbf{N} = \mathbf{T}' / \|\mathbf{T}'\|$.

From (33) we see that the derivatives of the tangent and principal normal vectors are not dependent on the orthonormal vectors (\mathbf{a}, \mathbf{b}) . It might be justified to use the term “curvatures” and symbols for the coefficients of the Frenet matrix (33):

$\chi \rightarrow \text{M-curvature}$

$\chi cm(\mathbf{T}, \mathbf{N}^\perp) := \chi_1 \rightarrow \text{M-second curvature},$

$\chi sm(\mathbf{T}, \mathbf{N}^\perp) := \chi_2 \rightarrow \text{M-third curvature.}$

With these symbols (35) can then be rewritten as

$$\begin{pmatrix} \mathbf{T}' \\ \mathbf{N}' \end{pmatrix} = \begin{pmatrix} 0 & \chi \\ -\chi_1 & -\chi_2 \end{pmatrix} \begin{pmatrix} \mathbf{T} \\ \mathbf{T}^\perp \end{pmatrix}. \quad (34)$$

Equation (34) (resp. (33)) corrects Equation (3.1) in [10].

Remark: Given three sufficiently differentiable or even analytic functions χ, χ_1, χ_2 one can calculate a (local) Taylor development of a curve γ which has these functions as curvature functions. Thus we can expect that these functions allow a formulation of a so-called main theorem, too.

7. SPACE CURVES IN MINKOWSKI 3-SPACE

In the following we give a short resume of the contents of [17], where the essential result is a description of the Frenet-Serret formulae in a Minkowski 3-space by curvature and torsion functions depending on a geometric parameter. For a space curve c we use a moving frame consisting of the unit vectors $(\mathbf{T}, \mathbf{N}, \mathbf{B})$ as usual, but with $\mathbf{T} \dashv \mathbf{N}, \mathbf{B} \dashv \mathbf{T}, \mathbf{B} \dashv \mathbf{N}$ in M_B^3 . Moreover, we use a so-called *deformation vector* \mathbf{d}_x and deformation function $\delta_x = \|\mathbf{d}_x\|$ of a unit vector \mathbf{x} in a two dimensional subspace of M_B^3 , which can be defined as follows:

$$\mathbf{d}_x := \mathbf{x}^{-\perp \perp} / \|\mathbf{x}^{-\perp \perp}\| - \mathbf{x} \quad (35)$$

Thereby $\mathbf{d}_x = 0$, if (and only if) \mathbf{x} and \mathbf{x}^\perp are mutually orthogonal.

Consider now a regular space curve c parameterized by $t \in [a, b]$. The tangents of c form a torsal ruled surface Φ . The generators of Φ are then given by the unit vector $\mathbf{T}(t)$. Moving this vector to the origin O of a fixed affine and M-orthonormal frame of a Minkowski 3-space M_B^3 , we receive the *M-spherical image* c , of Φ (and c) at the unit ball B . The cone $O \vee c_1$ is called the *direction cone* of Φ . Without loss

of generality we can assume that the parameter $t = s_1$ is the Minkowski arc length parameter of the curve c_1 , i.e. the derivation vector $\mathbf{T}'(s_1)$ is right-orthogonal to the vector $\mathbf{T}(s_1)$ all over the interval $[a, b]$ of s_1 . The curve c itself has an M-arc length parameter $s = s(s_1)$. After some calculations relating on the concept of deformation vectors we can get the Serret-Frenet matrix (see [17]) as follows:

$$\begin{pmatrix} \mathbf{T}' \\ \mathbf{N}' \\ \mathbf{B}' \end{pmatrix} = \begin{pmatrix} 0 & \chi & 0 \\ -\bar{\tau} & \bar{\tau} cm(\mathbf{N}, \mathbf{T}) + \bar{\tau}_1 cm(\mathbf{N}, \mathbf{B}) & -\bar{\tau}_1 \\ \tau_2 & -(\tau_1 + \tau_2 cm(\mathbf{N}, \mathbf{T})) & 0 \end{pmatrix} \begin{pmatrix} \mathbf{T} \\ \mathbf{N} \\ \mathbf{B} \end{pmatrix} \quad (36)$$

These Frenet equations are a special case of those for a skew ruled surface (see [17]), where we proposed the following concepts and notations for “curvatures” and “torsions”:

- i $\chi := \frac{ds_1}{ds}$... “M-curvature”,
- ii $\chi_1 := -cm(\mathbf{N}, \mathbf{B}')$.. 1st conical curvature,
- iii $\chi_2 := cm(\mathbf{T}, \mathbf{B}')$... 2nd conical curvature,
- iv $\chi_3 := cm(\tilde{\mathbf{B}}_N, \mathbf{N}')$... 3rd conical curvature,
- v $\chi_4 := cm(\tilde{\mathbf{T}}_N, \mathbf{N}')$... 4th conical curvature.

As M-torsion functions we declare

- i. $\tau_1 := \chi \chi_1$... M-torsion,
- ii. $\tau_2 := \chi \chi_2$... 2nd M-torsion
- iii. $\tau_3 := \chi \chi_3$... 3rd M-torsion
- iv. $\tau_4 := \chi \chi_4$... 4th M-torsion.

The functions $\bar{\tau}$ and $\bar{\tau}_1$ occurring in (36) depend on the 3rd and 4th M-torsion:

$$\bar{\tau} = \chi \frac{sm(\tilde{\mathbf{B}}_N, \mathbf{N}')}{sm(\tilde{\mathbf{B}}_N, \tilde{\mathbf{T}}_N) \|\tilde{\mathbf{T}}_N\|} \quad (37)$$

$$\bar{\tau}_1 = \chi \frac{sm(\tilde{\mathbf{T}}_N, \mathbf{N}')}{sm(\tilde{\mathbf{B}}_N, \tilde{\mathbf{T}}_N) \|\tilde{\mathbf{B}}_N\|} \quad (38)$$

Also here yields: Given sufficiently often differentiable M-curvature and M-torsion functions one can derive a (local) Taylor develop-

ment of c . For more details see [17].

CONCLUSIONS

We present a differential geometric approach to curves in Minkowski planes and spaces based on M-trigonometric functions, which turn out to have close relations to Brauner's M-angle measure. Throughout the paper we use Birkhoff-orthogonality in the Minkowski-spaces.

REFERENCES

- [1] J. Alonso, C. Benitez. Orthogonality in normed linear spaces, Part I, *Extracta Math.* 3(1) (1988), 1–15.
- [2] J. J. Alonso, C. Benitez. Orthogonality in normed linear spaces, Part II, *Extracta Math.* 4(3) (1989), 121–131.
- [3] G. Averkov, E. Makai, H. Martini, Characterizations of central symmetry of convex bodies in Minkowski spaces. *Studia Sci. Math. Hungar.* 46(2009), 493–514.
- [4] G. Birkhoff, Orthogonality in linear metric spaces, *Duke Math. J.* 1(1935), 169–172.
- [5] H. Brauner. Geometrie Projektiver Räume 1, 2. *B.I. Wissenschaftsverlag, Mannheim-Wien-Zürich* 1976. ISBN 3-411-01512-8.
- [6] H. Busemann. The Foundations of Minkowski geometry., *Comment Math. Helv.*, 24 (1950), 156–871.
- [7] M. M. Day, Some characterizations of inner product spaces, *Trans. Amer. Math. Soc.* 62(1947), 320–337.
- [8] S. S. Dragomir, J. J. Koliha, Two mappings related to semi-inner products and their applications in geometry, *Appl. Math.* 45(5), (2000), 337–335.
- [9] O. Giering. Vorlesungen über höhere Geometrie. *Vieweg, Braunschweig* 1982, ISBN 3-528-08492-8.
- [10] H. Guggenheimer. Pseudo-Minkowski Differential Geometry. *Ann. Mat. Pur. Appl.* 70 (1965), 305–370.
- [11] R. C. James, Orthogonality in normed linear spaces, *Duke Math. J.* 12(1945), 291–302.
- [12] R. C. James, Orthogonality and linear functionals in normed linear spaces, *Trans. Amer. Math. Soc.* 61(1947), 265–292.
- [13] R. C. James, Inner products in normed linear spaces, *Bull. Amer. Math. Soc.* 53(1947), 559–566.
- [14] H. Martini, K. Swanepoel, G. Weiss: The geometry of Minkowski spaces - a survey, *Part I, Expositiones Mathematicae* 19 (2001) 97–142, ISSN 0723-0869
- [15] H. Martini, K. J. Swanepoel, The geometry of Minkowski spaces-a survey. part II, *Expo. Math.* 22(2004), 93–144,
- [16] P. Milićić. On the B -Angle and g -Angle in normed spaces. *Journ. of Inequalities in pure a. applied math.* 8(2007) (3), Article 99, 9pp.
- [17] E. N. Shonoda, On Ruled Surfaces in three-dimensional Minkowski Space, *LAP LAMBERT Academic Publishing* (2011).
- [18] E. N. Shonoda, N. A. Omar, On Brauner's Angle and Frenet-Serret formulae in Minkowski space, *Sohag J. Math.* 1(2014), 1–7.
- [19] E. N. Shonoda, N. Omar. Some Special Curves in Normed Three Dimensional Space. *Appl. Math. Inf. Sci. Lett.* 2, 53(2), (2014), 53–57.
- [20] A. C. Thompson: Minkowski Geometry. *Cambridge Univ. Press* 1996, ISBN 0-521-40472-X

ABOUT THE AUTHORS

1. **Emad SHONODA** obtained his PhD in mathematical science at Institut für Geometrie, Technische Universität Dresden, Germany 2010. He is currently a mathematician in Faculty of Science, Port-Said University, Egypt. His e-mail address is en_shonoda@yahoo.de
Gunter WEISS, Dr. is a retired professor for Geometry (since 2011) at the Dresden University of Technology, Germany. His e-mail address is weissgunter@hotmail.com

BRUNELLESCHI'S DOME: AN OVERVIEW THROUGH 3D DIGITAL MODEL ABOUT GEOMETRICAL GENESIS AND PROPORTIONAL THEORIES

Silvia BERTACCHI
University of Florence, Italy

ABSTRACT: This research deals with the application of geometry and computer graphics as tools for documentation and dissemination of one of the most famous masterpieces of Italian Renaissance, the Brunelleschi's Dome, covering the church of Santa Maria del Fiore in Florence (Italy). By means of 3D digital models it is possible to deeply examine the geometric structure so as the constructive features that are actually linked to real building technique, allowing a comparison between the stated proportion by the architect and the achieved shape obtained when turning the design into the real world (with the consequent adjustment). This three-dimensional model, created by using solid modelling representation techniques, intends to be as close as possible to reality, although a simplification phase of metric and dimensional data has been necessary for two main reasons: on the one hand to manage the model, and on the other hand to underline, through an idealization process, the relevant geometric aspects of the original design. Idealized feature curves used for the creation of the solid model come from a highly detailed mesh of a 3D laser scanner survey carried out during the last years under the scientific supervision of the University of Florence. The complete digital design of the elements of the complex is essential to help both a general and detailed comprehension. It reveals the geometric genesis that leads and determines the final aesthetic-formal result, as well as the existing relations between the different parts and the whole, making real the object even in its hid-den components, not directly detectable to an outside glance but deeply connected to the structure, providing a less superficial view of relationships between geometry and the system inside the building. Moreover the model can also be used to become interactive, giving the possibility to query the linked information, e. g. about history, dimensions, original documents, or for 3D printing, in order to explain the different constructive steps of the building.

Keywords: Geometry, geometric analysis, dome, Florence, Brunelleschi, Italy.

1. INTRODUCTION

Since its construction, the Brunelleschi's Dome in Florence has been the subject of many researches on the geometrical construction adopted and on the analysis on its proportional system, in order to discover the building technique used to vault the huge eight-sided empty space left on the octagonal drum of the Florentine Cathedral [4, 6, 8, 10, 12, 21].

Most of the ancient studies on the famous masterpiece, whether founded on merely conceptual lucubrations or based on drawings and

surveys of the building and of its architectural elements, proved that the structure was subject to strict geometrical rules adopted by the architect.

In fact some of the most frequently used tools to control the architectural project and the final aesthetic outcome were drawing and geometry, together with scale models that made the formal control easier.

Nowadays we can definitely contribute to improve an in-depth knowledge of the dome by means of 3D digital models, representing an additional tool that researchers have at their

disposal for revealing architectural genesis in a more detailed way.

These models have different purposes – primarily a detailed documentation of the building and a simple dissemination via the Internet or by means of current real-time applications – and just as many ways to generate them with software from various data sources.

This contribution intends to illustrate some of the new possibilities of research on the geometrical genesis of the cathedral's dome, and new information gathered due to these detailed 3D digital models.

2. THE DOME OF SANTA MARIA DEL FIORE

Raised on the Duomo di Santa Maria del Fiore, the dome of Florence is one of the most famous and larger vaulted structures built during the Italian Renaissance (Figure 1).

Erected in less than 16 years, starting from 1420 until 1436, it completed at last the new church started in the previous century.



Figure 1: The Cathedral of Florence in the urban skyline. Photo by the author.

The earlier project for the enlargement of the pre-existing church of Santa Reparata had been designed by Arnolfo di Cambio at the end of the 13th century, when the city – increasingly important and prosperous because of trade – was in need of a wider religious building, representative of its splendour.

After more than a century the church was almost completed except for the dome. Unfortunately, the numerous changes the original design had undergone during this long period, made the construction of the vaulted structure so complicated that it seemed nearly impossible to achieve.

This was in a few words the premise to the crucial competition announced in 1418 for the construction of the dome, that represents the symbol of the city of Florence all over the world.

At that time Filippo Brunelleschi, who began his professional activity in the field of goldsmithry and clock-making, attracted attention also for his competence in architecture and engineering.

With his bright idea of constructing the dome without centering – thus saving money – he emerged as the winner of the competition and was nominated master builder, even though he was compelled to be assisted by Lorenzo Ghiberti, his bitter professional enemy.

The secret of his achievement was the peculiar building technique adopted, together with expedients for the placement of bricks (*corda blanda* [20], i.e. curved lying beds to avoid discontinuities and weak static areas especially in correspondence with the corners and *spinapesce* [15], i. e. bricks laid with the herringbone technique usually utilized in the construction of Byzantine self-supporting domes), that allowed the construction without expensive wooden supporting falsework.

In this way the architect could complete the structure made of a double shell (the internal with a static function while the external for protection against atmospheric agents and aesthetic reasons) and sustained by the octagonal

irregular drum, at a height of about 55 meters from the base of the church, and following the geometrical rules of particular pointed arches.

The internal diameter of the drum from corner to corner reached 77 *braccia fiorentine* (= 45 m), a local unit of measurement (arm's length of Florence) equal to almost 0.584 meters [16], an exceptional dimension at that time; from side to side 72 *braccia*; the external diameter was about 92 *braccia* (= 54 m), each measure proportionate to the main elements of the church. The eight vaulted sections were built with the total height of 116 meters including the marble lantern finished after Brunelleschi's death, being each one about 36 meters high (Figure 2).



Figure 2: Brunelleschi's Dome from the bell tower known as *Campanile di Giotto*.
Photo by the author.

2.1 Ancient models of the dome

The study of architecture and its elements by examining models is a common practice from time immemorial.

In ancient times the use of architectural

models was the only three-dimensional tool – together with drawings and sketches – to represent a complex architectural project and it was used by engineers, architects and artists in order to better control the formal final outcome [18].

Thus made of wood or wax, from plaster or clay, the two main purposes of the small-scale models were on the one hand to show a preview of the artistic creation to the patron, the public purchaser or non-experts, and on the other hand they were used to check the project accuracy and its static structure.

According to some scholars the model is not only an isomorphic miniature to study the static features, not often with the same performances as a full-scale building (Vitruvius distrusted using the scale model as a copy of the building, due to the differences in the physical behaviour), but it has a symbolic value as the best representation ever of the designed architecture what is to be [22]. The geometrical model is an ideal copy of reality, in full-scale, obtained thanks to the synthesis by an operator and consists of pure geometrical shapes that are similar - but not identical due to necessary simplifications - to the irregular structure of the real object. This model can be reduced to scale with an approximation operation, and then with a geometrical operation of projection and section the object can be represented on the drawing sheet to document its shape [11, 17].

All things considered, in fifteenth-century Florence the practice of building models was common, especially concerning the design of the cathedral's dome.

Some of these wooden models, traditionally attributed to Filippo Brunelleschi, have survived to this day: the first one represents the dome and the tribunes (scale 1:50) and according to the scholars it should be coeval to the construction (Figure 3) [13].

The second is the representation of the marble lantern, dating back to 1436, when Brunelleschi won the competition for the project [19].

Similarly, the models for the decoration of

the structure of the dome with balconies, built in 1507 for the new competition, are still preserved at the Museo dell'Opera di Santa Maria del Fiore.



Figure 3: Wooden model of the dome and tribunes of the cathedral attributed to Filippo Brunelleschi, dating back to approximately 1420-1440, Museo dell'Opera di Santa Maria del Fiore, Florence.

Concerning the dome, it seems that Brunelleschi had built some models in secret, or never completely explicit, being afraid of giving information to his rivals about constructive solutions. According to historical reports Brunelleschi had also drawn a full-scale section on a large flat space on the banks of the Arno River, in order to give a detailed outline of the curvature of the dome and to control the geometrical features of the ribs [7].

The only masonry model of the dome had been built in the vicinity of the cathedral after the competition announced in 1418 as a guideline for the construction of the dome, replacing the previous one dating back to 1368, and was destroyed in 1431 nearly at the end of the con-

struction works. Its dimensions (2 x 4 m) were large enough to enter into the model and get an impression of the project.

3. 3D DIGITAL MODELS

In the later years of the Twentieth Century, thanks to the technological progress and the development of innovative computerised modelling tools, the scholar's toolkit has been considerably enriched in respect to the recent past.

Presently we can create three-dimensional digital models of different kinds, with a variable level of detail and for a multiplicity of intents [14].

Generally speaking, unlike analogue models, digital ones are really more adaptable, concerning the possibility to undergo modifications, to be more easily portable and disseminated, in addition to the advantage of real-time visualization that makes the interaction with the virtual space easier also for non-experts.

In any way, regardless of the software used, the construction of the computerized model is not automatic but the result of the knowledge and skill of the designer and the critical use of digital technologies.

As previously said the creation of digital models can be based on different data sources, gathered from reality or idealized, that is to say based on 3D point cloud and mesh model or on conceptual models based on the ideal geometric genesis.

In addition, there are more and more tools for the data management and handling.

For instance, the reverse modelling tools are specifically designed for the conversion of 3D meshes from active and passive sensors into mathematical models (NURBS or solids). Their use is widely known and developed in the field of design, mechanic and civil engineering since these application areas need a strict control of the 3D model accuracy.

In the field of Cultural Heritage (archaeological remains, historical buildings), the three-dimensional mesh is preferable to mathematical models for dissemination practices,

since it provides efficaciously an accurate texturing, optimization and light simulation rather than achieving conceptual shapes through parametric patches.

In other words, reverse modelling applications are mainly employed for point cloud triangulation, mesh healing, big files decimation, segmentation of complex models into smaller elements for an easier use and so on, but their more effective interactive tools for slicing the model (and getting sections or feature curves and surfaces) are often underused, although very suitable when concerning shape analysis.

As a matter of fact, highly detailed meshes are more effective when an accurate representation of the system is required, even though irregular, e.g. for structural deformations in order to quantify the cyclical movements of the cracks and fractures, detachment and loss of material from the surface, thermal fluctuations due to weather conditions.

On the contrary NURBS mathematical models' strong point is the possibility to model continuous smooth surfaces, with nice curvature flow.

In short, on the one hand there are mesh models, easy to manage into a multiplicity of applications, ranging from entertainment software to reverse modelling applications for data extraction (sections and projections in this particular case).

On the other hand reverse modelling applications are considered more fitting with redesign process, which main outcome are NURBS or solid representations of scanned objects.

Frequently the majority of operators simply extracts poly-lines and splines from point cloud or highly detailed meshes and then use them into computer-aided design applications, generating 3D digital models by loft, sweep, extrusions, and curve networks tools.

Reverse modelling applications are indeed the more performing and all-embracing solution to explore and simplify curves and patches obtained from high-poly models, since they provide several aspects for a comprehensive and

reliable exploration of a shape:

- 1) Tools for the creation of reference planes, vectors and points on the base of massive data quantities (e. g. cutting planes);

- 2) Tools for extraction of poly-lines and splines: using reference geometries or their instances in an interactive way;

- 3) Tools for extracting 2D and 3D primitives (circles, ellipses, lines, spheres, planes, etc.) using high-poly meshes as reference templates;

- 4) Tools for the control and measuring of section deviation between point-clouds, high-poly meshes, NURBS surfaces.

For this reason it is important to point out that software developers in the framework of reverse modelling applications implemented both documentation and design tools.

Starting from this data by means of solid modelling it is thus possible to create conceptual models – necessarily reduced and simplified for an easy management – approximating the existing surfaces of the real *vela's* intrados.

In the present paper, referring to the Florentine dome, pros and cons will be pointed out, comparing the ample scopes for documentation and dissemination of a critical use of both of them.

3.1 Mesh model from laser scanner data

In recent past there was no possibility to achieve a metric survey with a millimetric level of detail, to the purpose of precisely showing the exact spatial shape of the dome.

Latest digital devices allow carrying out metric surveys with an increased reliability, especially thanks to developments of laser scanner technologies and their applications in the architectural and archaeological field [5].

It is thus possible to update the outdated surveys of the dome, have at researcher's disposal more reliable measurements than in the past and achieve accurate polygonal mesh models of existing architecture as close as possible to the real building.

In fact this kind of measuring, among the many advantages, allows a better accuracy of

data capture; the following step consists in converting the point cloud into a detailed mesh representing the irregularity of the real architectural element.

In this sense, this information can be useful from the point of view of static surveys for an overall systematic description of the cracks occurred in the dome and moreover for the possibility of up-to-date analysis regarding the surfaces of the vaulted sections (for example a monitoring on seasonal deformations).

Regarding the dome of Florence, the most recent surveying activities carried out dates back to 2004 (A.B.C. General Engineering in collaboration with Codevintec) and 2006 (Geoarte S.T.A.), under the guidance of the University of Florence, with two laser scanner campaigns by means of time-of-flight devices to gather the entire intrados and the north-eastern area of the extrados of the double shell [3] (Figure 4).

Thanks to this point cloud of the intrados from the survey it has been possible to create a reliable polygonal mesh model (with INUS Technology Rapidform XOR3), aimed at studying the proportional structure and drawing new conclusions on its geometrical system.

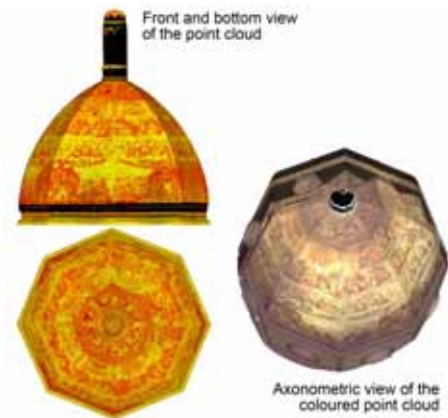


Figure 4: 3D laser scanner survey of the interior of the Brunelleschi's dome, Geoarte S.T.A., surveying campaign year 2006.

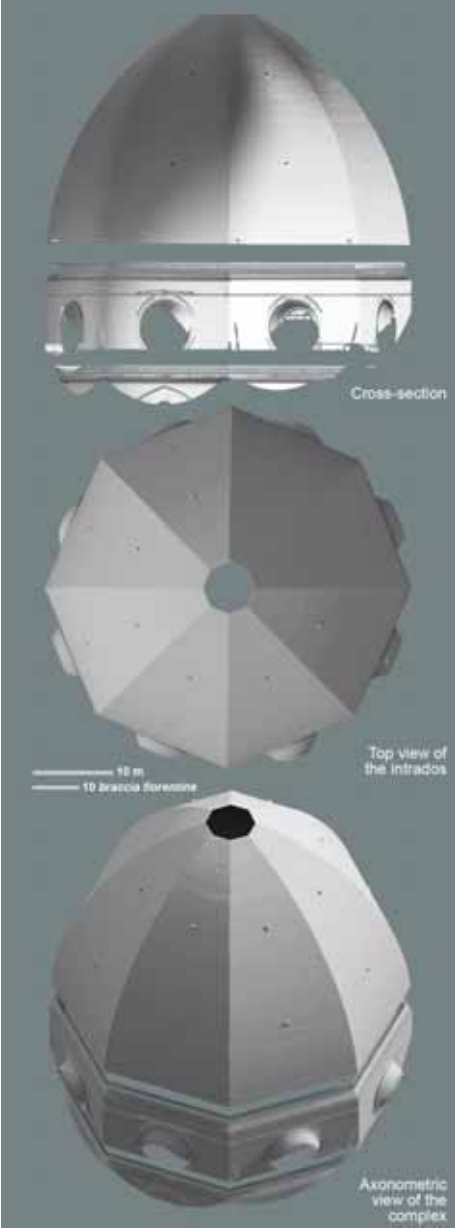


Figure 5: Orthographic and axonometric view of the mesh model of the dome's intrados.

The model is interactive and any cutting plane can generate a section (Figure 6).

But first of all, it is fundamental to treat defects due to the conversion and correct the topological errors through a resampling, in order to have an isotropic mesh within the tolerance limits of the scanner accuracy. Then it will be possible to choose the cutting elements and their position, e.g. reference planes, vectors or reference points, for the purpose of generating sections (Figure 7).

These ones can be divided mainly into two different kinds, depending on the future use.

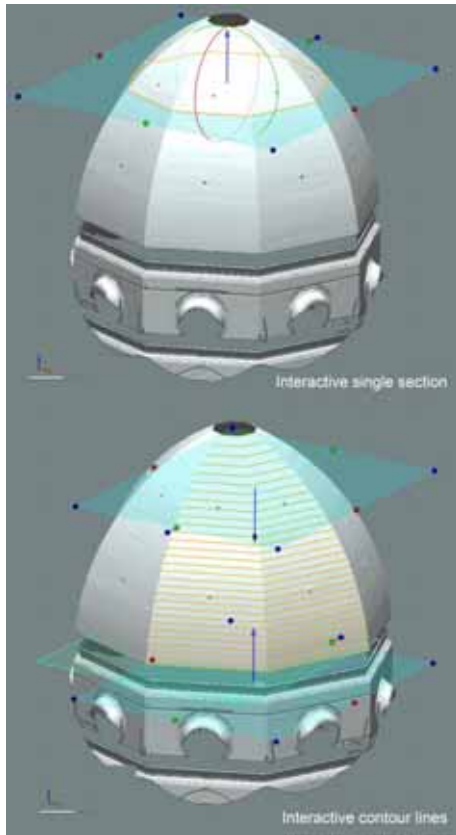


Figure 6: Interactive tools for the determination of sections and contour lines on the model.

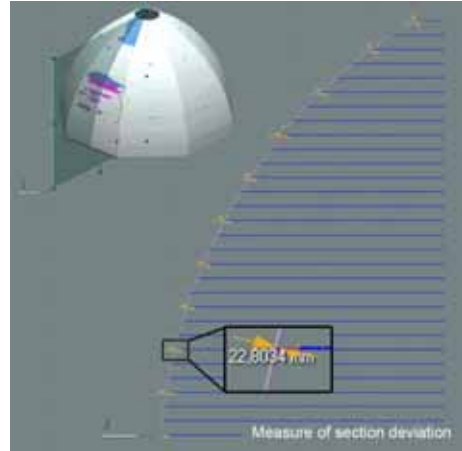


Figure 7: Measure in millimetres of sectional difference between mesh model (grey) and surface model (violet); detail of the gap.

The first sections – polylines or curves – are used for documentation scope, since they represent a real object in a reliable way (they are composed of a large number of vertices), but they are not suitable for modelling. The second kind of sections is useful to re-draw directly on the mesh using the latter as reference.

In this software there are CAD style commands, like lines, circles, ellipses and all the primitive elements: all these geometrical shapes are drawn using reference sections obtained from mesh models.

Even if reverse modelling commands are addressed to mechanical engineering and product design, they can also be suitable for Cultural Heritage representation, first and foremost because the research of shapes anchored to the mesh is simplified.

In fact, the automatic interpolation makes easier and faster the research of best fitting circles on the interior corners of the dome (as shown in the examples of Figure 8) and all the sections are already positioned in a 3D environment.

Some scholars recently [9] achieved interesting results with the proprietary software, an-

analysing the reference sections obtained by the laser scanner survey, but considering only the corner section and the one in the middle of the octagonal vaulted sector as the most interesting points to check.

On the contrary, thanks to these tools for 3D drawing it is even more convenient to consider the geometry in its entirety, because of the accurate control of a multitude of spatial information. As an instance by means of the interactive selection it is possible to do an estimate of the deformation events of such a big dome.

Furthermore, the primitive curves obtained can be directly used as a template for modelling (especially in NURBS and solid modelling applications).

3.2 Solid conceptual model

The method used for the creation of the solid model of the intrados is based on several steps and by adopting different principles. The modelling was generated using as template the best fitting curves, obtained as described before by the polygonal mesh by the survey.

Contour lines were used to determine irregularities on the wide surfaces of the vaulted sectors, certainly not appreciable considering only 2 or 3 reference sections.

The results of these operations had to be interpreted through the geometrical knowledge, considering the adherence of reality to the requested proportion ("*sesto di quinto acuto*" = rule of the *pointed fifth*) established even before Brunelleschi's participation to the construction, and in the light of the historical period and available building techniques, as it was impossible to build a perfectly symmetrical dome.

The outside part of the model instead was generated using traditional or integrated surveys (archival material, photogrammetric surveys), checked for their reliability and compared in [9].

The complete model reveals the interior structure of the dome as well as the actual connections among the architectural elements and the whole complex (Figure 9).

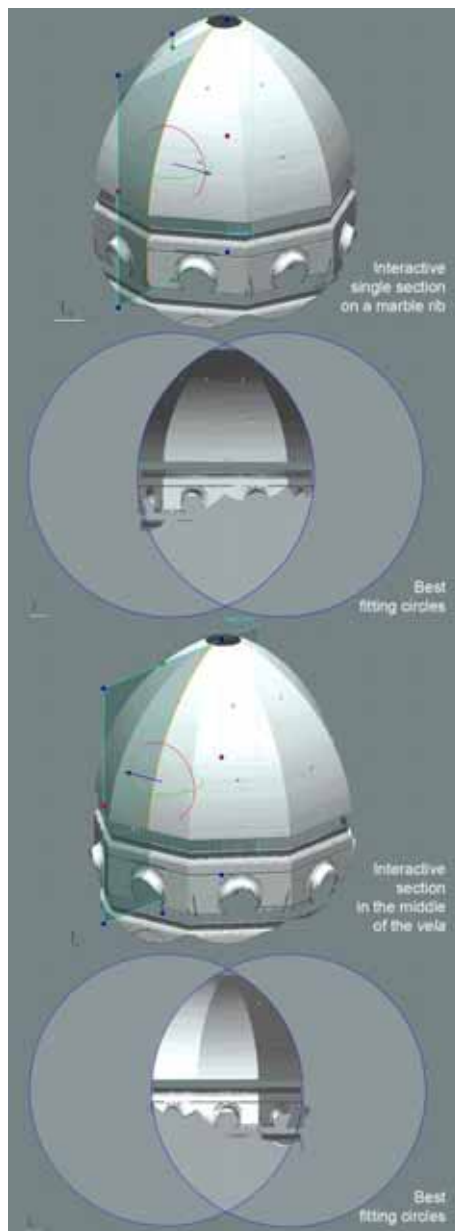


Figure 8: Different sections and related best fitting circles by automatic interpolation.

All the structural components follow a semantic partitioning.

For example the elements of horizontal connection as the walkways between the two shells, stone and wooden chains, the horizontal arches between the buttresses, or the vertical ones, such as the marble ribs and the corner spurs.

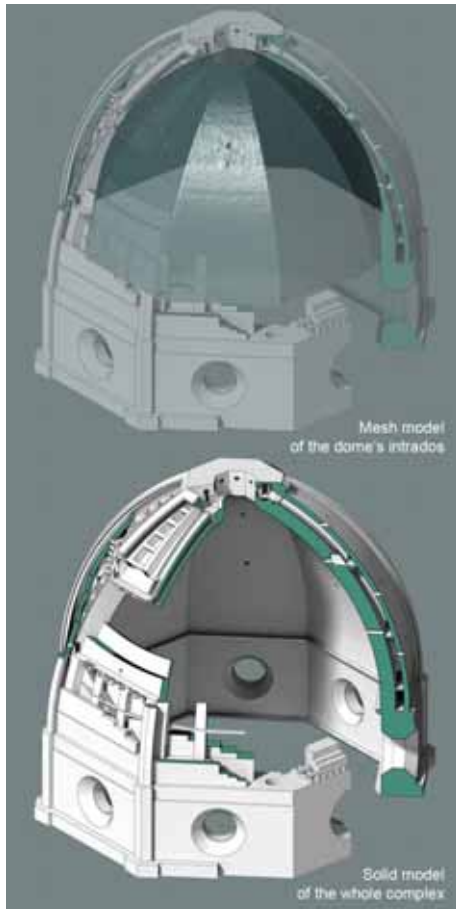


Figure 9: Modelling process starting from the mesh model (transparent colour) of the intrados by the 3D survey; bottom: the solid model obtained, showing the elements.

4. GEOMETRICAL GENESIS AND PROPORTIONAL THEORIES

In accordance with tradition, the dome's silhouette should have a pre-established shape, not undergoing any change from the project to the real construction.

For the eight sections of the dome it was requested a pointed arch shape, in the inner profile following the rule of the pointed fifth (*"sesto di quinto acuto"*), being the centre of the circumference drawing the corner's outline at a fifth of the distance between two opposite vertex of the octagonal drum. The external shell had to be more swollen, adopting the *pointed fourth* rule (*"sesto di quarto acuto"*), with the centre moved to a quarter of the distance (Figure 10).

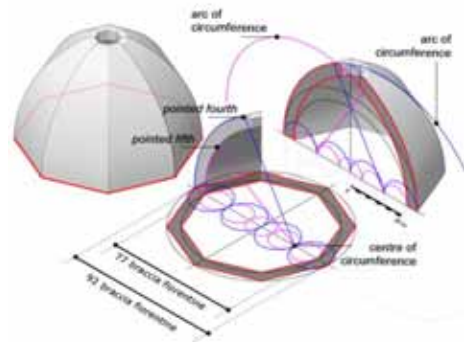


Figure 10: Proportions of the theoretical model based on the rules of the *pointed fifth* and the *pointed fourth* [2].

This matter lied at the hearth of Florentine citizens, insomuch as they were concerned about the vague project by Brunelleschi and when the building was still under construction, the architect was criticized harshly for the interior and exterior outlines and accused not to respect the requirements.

In the renowned parchment by Giovanni da Gherardo da Prato, dating back to 1425-26, the hand-drawn sketch showed a clear geometrical layout fundamental to build correctly the dome.

From that moment on, numerous are the hypothesis made on this issue [9, 12, 21].

The most recent studies on the analysis of the geometrical genesis are based on sections gathered from the 3D point cloud from the digital surveying campaigns [9].

According to this study it seems that actually the dome geometrical system is based on circumferences and ellipses.

In particular, the genesis of the eight sections starts from the drawing of an arc of circumference on the diagonal side of the base octagon, following specific rules. In this way the section on the corner is represented by a portion of circumference by construction.

On the contrary, when sectioning in the middle of the *vela* a part of an elliptical arc is obtained.

Due to this assumption, the surfaces of the eight sides are determined by horizontal lines (generators of the cylinder) that move along an elliptical path, visible in the mid-section of the side, and in theory the *vela* should overlap on an elliptical cylinder (Figure 11).

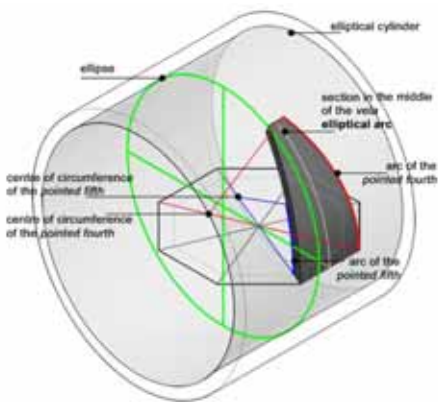


Figure 11: Geometric construction of the elliptical cylinder starting from the circular arc of the corner profile [2].

But when using deviation tools of reverse modelling software, it is evident that the over-

lapping of the conceptual model on the point cloud cannot be perfect, as shows the colour range and schematic table (Figure 12).

In the image there is a summary of the procedure to calculate the deviation measure between the highly detailed mesh resulting from the point cloud and the idealized model.

Best fitting curves are used for the generation of a NURBS patch, and then compared to the original mesh from scanner.

The surface is mapped with a colour range, in blue for negative deviation and red for positive one. The colour key, for the extension on the whole surface of a sector of the dome, shows very clearly the lack of uniformity of the deviation. This kind of visualisation allows a better understanding of the 3D models of such a huge object.

Non-linear and unexpected displacements are underlined by different colours.

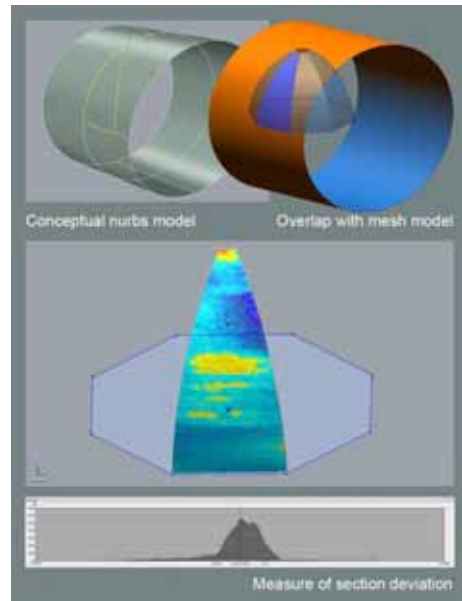


Figure 12: Measure of the section deviation between the conceptual mathematical model and the numerical model of the intrados.

In case of a conceptual model it is important to establish a bias, a threshold between a simplified/idealized model of a real architecture and its fully detailed representation. In this case the average deviation is equal to -15.495 mm, while the 79.42% of the samples presents errors ranging from -32.10 and +1.11 mm for the specific example of the *vela*'s sector. Similar results have been achieved through the application of this procedure on the other sectors.

The construction of the dome's model is then a complex matter, due to the different nature of input data, to its morphological detail, and last but not least to the planning and management of the modelling process itself: complex shapes, including the exterior and the interior ones, as well as the structures of the dome needed to undergo a difficult segmentation phase aimed at organising in a hierarchy all the parts forming the model [1].

5. CONCLUSIONS

In conclusion, it is evident that the updated technologies, as well as latest devices, can both help the geometrical research in an increased reliability, even on topics fully documented by many scholars.

In particular, dealing with geometrical analysis and having data available from a laser scanner survey, it is well-recommended to study the monument by means of reverse modelling software, which tools provide a full set of commands for measuring deviation, both in case of mesh resampling and editing or shape analysis.

In the final analysis one can say that nowadays it is absolutely impossible, as well as methodologically mistaken, not to use up-to-date technologies for the in depth and critical research on architectural heritage and its proportional and metric dimensions.

These digital models, as previously demonstrated, have a scientific purpose and they are not only a mere and optional virtual representation, owing to the encoding in the model itself of different fields of knowledge.

ACKNOWLEDGMENTS

Acknowledgments to Professor Roberto Corazzi, University of Florence, for the laser scanner data of the dome, captured during the surveying campaign by the company Geoarte S.T.A. (year 2006).

REFERENCES

- [1] B. Benedetti, M. Gaiani, and F. Remondino. *Modelli digitali 3D in archeologia. Il caso di Pompei*. Edizioni della Normale, Pisa, 2010.
- [2] S. Bertacchi. The dome of Santa Maria del Fiore in Florence: a virtual model. In G. Tampone, R. Corazzi, and E. Mandelli (edited by). *Domes in the world, Congress Proceedings, Florence, March 19-23, 2012*, pages 1-10. Nardini editore, Firenze, 2012.
- [3] M. Bini and C. Battini (edited by). *Nuove immagini di monumenti fiorentini. Rilievi con tecnologia scanner laser 3D*, pages 50-51. Alinea, Firenze, 2007.
- [4] D. Cardini. *Il bel San Giovanni e Santa Maria del Fiore: il centro religioso di Firenze dal Tardo Antico al Rinascimento*. Le Lettere, Firenze, 1996.
- [5] E. Chiavoni and M. Filippa (edited by). *Metodologie integrate per il rilievo, il disegno, la modellazione dell'architettura e della città: ricerca Prin 2007, coordinatore nazionale Mario Docci*. Gangemi, Roma, 2011.
- [6] C. Conforti. *Lo specchio del cielo: forme significati tecniche e funzioni della cupola dal Pantheon al Novecento*. Electa, Milano, 1997.
- [7] R. Corazzi. *La Cupola del Brunelleschi, Il segreto della costruzione*. Angelo Pontecorboli Editore, Firenze, 2013.
- [8] R. Corazzi and G. Conti. *La cupola di Santa Maria del Fiore raccontata dal suo progettista Filippo Brunelleschi*. Sillabe, Livorno, 2005.

- [9] R. Corazzi and G. Conti. *Il segreto della Cupola del Brunelleschi a Firenze. The Secret of Brunelleschi's Dome in Florence*. Angelo Pontecorboli Editore, Firenze, 2011.
- [10] R. Corazzi, G. Conti, and S. Marini. *Cupola di Santa Maria del Fiore: tra ipotesi e realtà: studi e ricerche per un'indagine avanzata*. Pitagora, Bologna, 2005.
- [11] L. De Carlo (edited by). *Informatica e fondamenti scientifici della rappresentazione. Seminario/laboratorio, Roma 12-14 febbraio 2007, Strumenti del dottorato di ricerca in scienze della rappresentazione e del rilievo*. Nuova serie, vol. 1. Gangemi, Roma, 2007.
- [12] S. Di Pasquale. *Brunelleschi: la costruzione della cupola di Santa Maria del Fiore*. Marsilio, Venezia, 2002.
- [13] G. Fanelli and M. Fanelli. 2004. *La Cupola del Brunelleschi. Storia e futuro di una grande struttura*. Firenze: Mandragora, 2004.
- [14] F. Fantini. Modelos con nivel de detalle variable realizados mediante un levantamiento digital aplicados a la arqueología. In *EGA expresión gráfica arquitectónica*, 19: 306-317, 2012.
- [15] F. Gurrieri and C. Acidini Luchinat. *La cattedrale di Santa Maria del Fiore. Volume primo*. Cassa di Risparmio di Firenze, Firenze, 1994.
- [16] L. Ippolito and C. Peroni. *La cupola di Santa Maria del Fiore*. Carocci, Roma, 2007.
- [17] R. Migliari. *Geometria dei modelli. Rappresentazione grafica e informatica per l'architettura e per il design*. Edizioni Kappa, Roma, 2003.
- [18] H. A. Millon. I modelli architettonici nel Rinascimento. In H. A. Millon and V. Magnago Lampugnani. *Il Rinascimento da Brunelleschi a Michelangelo, la rappresentazione dell'architettura*, pages 19-73. Bompiani, Milano, 1994.
- [19] G. Morozzi. *L'Architettura fiorentina nei plastici antichi e moderni*. Gruppo Amici della Certosa di Firenze, Firenze, 1985.
- [20] P. A. Rossi. *Le cupole del Brunelleschi: capire per conservare*. Calderini, Bologna, 1982.
- [21] P. Sanpaolesi. *La cupola di Santa Maria del Fiore: il progetto, la costruzione*. EDAM, Firenze, 1977.
- [22] M. Scolari. *Il disegno obliquo. Una storia dell'antiprospectiva*, pages 131-189. Marsilio, Venezia, 2005.

ABOUT THE AUTHOR

1. Silvia Bertacchi is architect and European Ph.D. in Science of Surveying and Representation of Architecture and Environment at the Department of Architecture of the University of Florence (Italy). She is tutor of "Drawing" for the Bachelor Degree in Construction Engineering at the Alma Mater Studiorum, University of Bologna, Campus of Ravenna (Italy) and visiting assistant professor of "Descriptive Geometry" for the Master Degree in Architecture at the University of Ferrara (Italy). Her research is mainly focused on Descriptive Geometry and its application on architecture. Since 2010 for her PhD thesis she has been studying the compositional models in the field of military architecture dating back to the 16th century, deeply analysing the geometrical tool used for defensive buildings.

CAUSTIC SURFACE AND QUASIFOCAI LINE

Alexander DVORETSKY and Tetiana DENYSOVA

Academy of Environmental Protection and Resort Development,
Crimea, Russian Federation

Abstract. Of particular importance for the practice of lighting and solar technology is to define the parameters of the zone of greatest rays concentration reflected by curved mirror and the determination of the reflected flux distribution on the surface of a receiver.

Curves of caustics by reflection for 2D and surfaces of caustics by reflection for 3D are considered. It is proved that the surface of caustics of the reflected ray congruence is not the zone of maximum concentration. Such zone is quasifocal line. For 3D the curve of caustic can be considered as the zone of concentration of reflected rays, so called “burning curve”.

Keywords: reflecting surface, caustic surface, quasifocal lines, “burning curve”, reflected flux distribution, energy map.

1. INTRODUCTION

Reflections play a central role in many applications: the conversion of solar energy into other forms of energy, lighting natural and artificial light in architecture and so on.

Particular importance for practice is determining the parameters of the greatest concentration zone of reflected rays congruence. There is a perception that this area is caustic: flat and spatial. However, there is a reason to assume that the area of greatest concentration of the reflected rays congruence is quasifocal line [4, 6].

Two-dimensional caustic as the envelope of the reflected rays (2D) belongs to the plane, if the normals of the planar section curve coincides with the normals of the surface along this section.

To solve three-dimensional (3D) reflection task the surface of normals along the planar section of the reflecting surface and the reflected rays surface along the same section is proposed to consider [9].

The envelope of the normals of a curve is known as its evolute (and also be thought of as the set of centers of curvature of the curve). Thus focusing occurs on the evolute of W .

Because of this optical connection, the evolute is also known as the focal set of W , or indeed the caustic by reflection. According to Brocard and Lemoine [1], the orthotomic was introduced into the study of caustics by Quetelet in 1826. Caustics themselves to go back as far as Tschirnhaus (1682), a detailed study of special cases was carried out by Cayley [2] in 1856.

In the works of Bruce, Giblin and Gibson [5] the caustics and orthotomics with point source, located not in the ellipse focus are considered. Figure 1a shows the reflection by an ellipse with the light source located near a focus of an ellipse. Figure 1b shows reflected from the light source located further from the focus; the caustic has split into two. In the latter case, the caustic has two cusps, orthotomic has one self-intersection point.

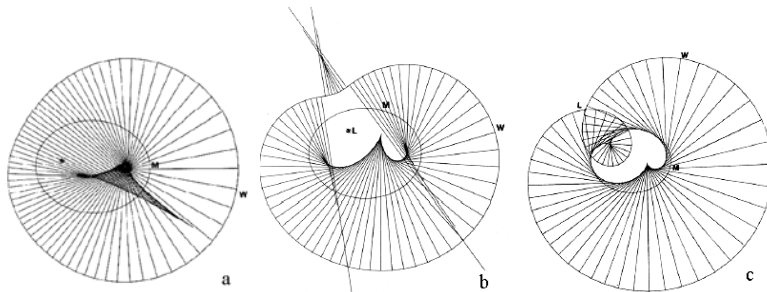


Figure 1: Reflection by an ellipse with the different location of the light source.

Figure 2 is a computer drawing of part of a caustic for an ellipsoid (produced by David Fidal). Caustic is the envelope of the projection of the reflected rays for the cross section of the ellipsoid.

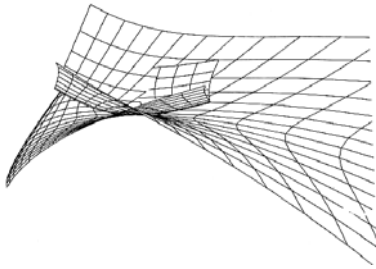


Figure 2: Caustic for an ellipsoid.

It is known that a caustic of a circle with a point source in infinity is epicycloid (Fig. 3). Its equation is

$$\begin{aligned} x &= \frac{3a}{4} \cos \varphi - \frac{a}{4} \cos 3\varphi, \\ y &= \frac{3a}{4} \sin \varphi - \frac{a}{4} \sin 3\varphi, \end{aligned} \quad (1)$$

where a – radius of a reflecting circle,
 α – angle, that determines the point on the reflecting circle.
 It is a "burning curve".

In the article [3] professor G. Glaeser considered a caustic of a circle for different positions of the source - planar task and surface - spatial tasks (Fig. 3). For the planar task there is fair statement that the caustic is "burning curve".

In the case of spatial task all the reflected rays congruence tangent to the surface, which is a caustic by reflection or a focal surface.

Caustic of a reflecting sphere with a point source in infinity is the rotation surface of an epicycloid (Fig. 4). In this regard, the question arises: - "What is "burning zone" for the reflected rays congruence: a caustic surface or a quasifocal line?"

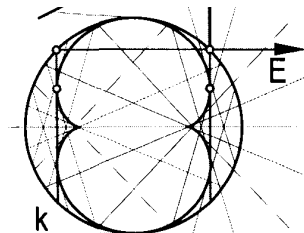


Figure 3: Planar caustic for the circle.

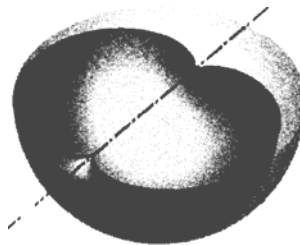


Figure 4: Caustic surface of the sphere.

2. QUASIFOCAL POINT

The quasifocal point F for planar task is presented in Figure 5. The reflecting curve is a circle. The spatial visualizations of rays reflected by a sphere and the quasifocal point F that locates in the middle of the radius is shown in the Figure 6.

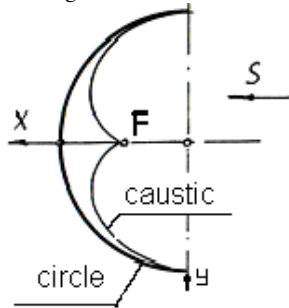


Figure 5: Quasifocal point F (planar task).

The reflected rays surface [3, 4] with a ray source in infinity, as a ruled surface, is set by three director lines. In an article [6, 11] was proved, that such director lines are the following: flat section k along which the reflected rays are under construction; a double line m , in which the rays are in pairs crossed; a director cone T of the reflected rays surface Ω .

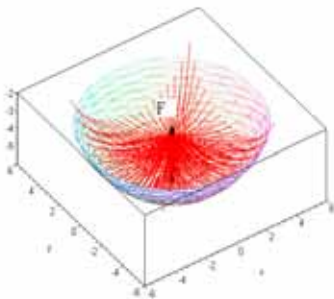


Figure 6: The spatial visualizations of rays reflected by sphere (spatial task).

The axonometric of the reflected rays surface Ω along a circle k is represented in Figure 7. A normal surface along a director k of a canal surface or surface of rotation is the direct circular cone with top in a point T . The incident rays are parallel to the direction S .

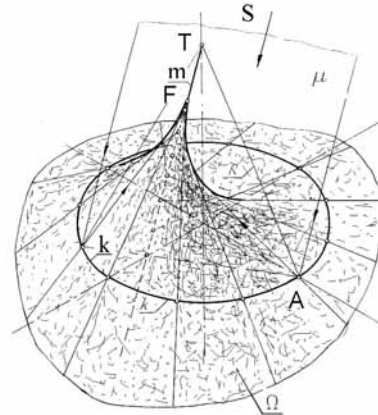


Figure 7: Reflected rays surface.

In this connection, it is possible to make a conclusion, that the basic share of a solar energy reflected from a half generator k , is concentrated near to the point F . In the further reasoning, with accuracy, necessary for practice, we shall consider the point F as concentration of the reflected rays.

Definition. Quasifocal point F is one of the two boundary points on a double line of the reflected rays surface, near to which the basic share of a reflected rays are concentrated, reflected by a part of a generator of the concentrating surface, corresponding to the aperture angle.

2. QUASIFOCAL LINE

In the 2D reflection problem, there is the concept of "a burning curve". In this case, "a burning curve" is the caustic by reflection. For the 3D problem, caustic by reflection is a surface [6]. The objective of the report is to prove the fact that the caustic surface is not the focal zone of reflected rays. In fact, for the 3D task, the area of maximum concentration is the quasifocal line (Fig.8).

Quasifocal lines and caustic surfaces for several reflecting surfaces such as surfaces of revolution and canal surfaces are studied.

This theoretical result is useful to geometrical simulation of the receiver surface, reflecting surface, reflected flux distribution and so on.

2.1. Quasifocal line for a surface of revolution.

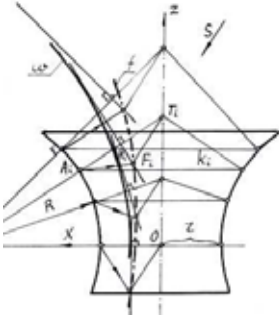


Figure 8: The quasifocal line \mathbf{f} that formed by rays, reflected by the surface of rotation with circular generator.

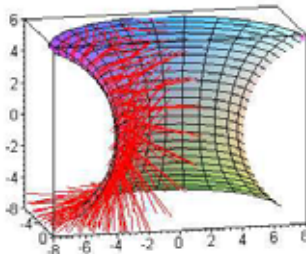


Figure 9: The flux reflected by the surface of rotation.

Definition. All quasifocal points of one-parametrical set of the reflected rays surfaces form **quasifocal line \mathbf{f}** [8] that corresponding to the reflecting surface.

Parametric equation of the quasifocal line \mathbf{f}

$$x = \frac{ctgt [(R[\cos t] + 1) + R[\sin t] tgt] d}{2(\frac{p}{m} + ctgt)},$$

$$z = x \frac{p}{m} + tgt \quad (2)$$

The spatial visualizations of quasifocal line \mathbf{f} for the surface of rotation are shown in Figure 9.

The reflecting surface of rotation with circular generator and the construction of a quasifocal line were considered in article [8].

2.2. Quasifocal line for a cone.

Statement. The quasifocal line that formed by rays, reflected by the direct circular cone is a straight line \mathbf{f} , taking place through the top of the cone and is seen in Figure 10. The inclination angle of the quasifocal straight line depends on the angle φ at the top the reflecting cone and the direction of incident rays S (m, n, p).

The equation of a direct circular cone (Figure 10) in system xOz is

$$\frac{x^2}{tg^2 \varphi} + \frac{y^2}{tg^2 \varphi} - \frac{z^2}{1} = 0. \quad (3)$$

The quasifocal point F will be defined

$$z_F = \frac{m[\sin 2\varphi + c[p \sin^2 \varphi (tg^2 \varphi \cos^2 \varphi)]}{m[\sin 2\varphi + 2p \sin^2 \varphi]},$$

$$x_F = \frac{m}{p}(z_F \cos \varphi). \quad (4)$$

The equation of the quasifocal line that passes through the origin of coordinates looks like

$$x = \frac{x_F}{z_F} z. \quad (5)$$

For the cone, the quasifocal line \mathbf{f} is straight line.

Angle of an inclination of the quasifocal line \mathbf{f} is

$$\angle = \arctg \angle, \text{ where } \tan \angle = \frac{x_F}{z_F}. \quad (6)$$

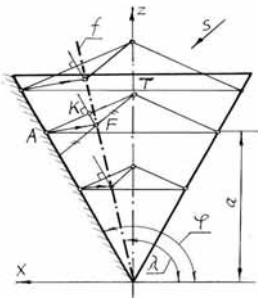


Figure 10: The quasifocal line \mathbf{f} for the direct circular cone.

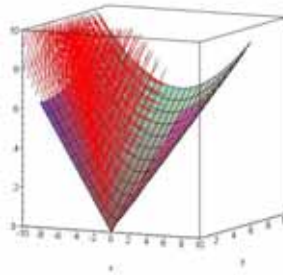


Figure 11: The reflected flux by reflection of the cone.

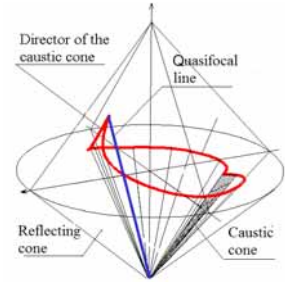


Figure 12: The caustic surface by reflection of the cone.

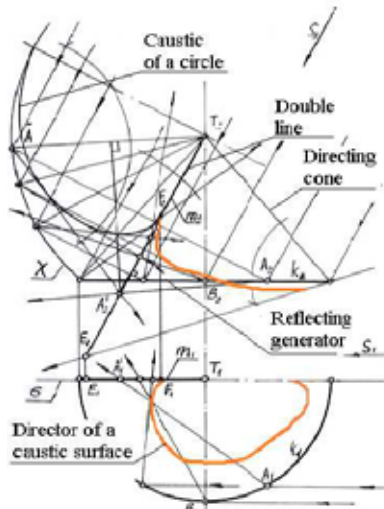


Figure 13: Director of a caustic surface.

As a result of computer simulation the caustic surface of a paraboloid of revolution under the direction of the incident rays at 30° to the axis of the paraboloid is obtained (Fig.14) [10].

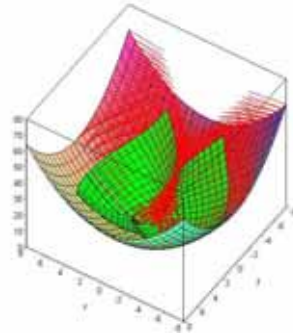


Figure 14: The caustic surface by reflection of the paraboloid.

2.3. Quasifocal line for torus reflector.

The equation of the quasifocal line \mathbf{f} for the torus with the different angles of incident rays

$$\begin{aligned} x &= \frac{r[\cos(t)[\cos(t) + \sin(t)]g(t)]\sin(\arctg(\frac{tg(t)}{\cos(t)})]}{2[\cos(t)g(t)]\sqrt{1+tg^2(t)}} + R\cos(t), \\ y &= R\sin(t), \\ z &= \frac{r[\cos(t) + \sin(t)]g(t)]\sin(\arctg(\frac{tg(t)}{\cos(t)})]}{2\sqrt{1+tg^2(t)}}, \end{aligned} \quad (7)$$

where α – the angle between an incident ray and the plane xOy .

The model of the torus reflector, the flow of reflected rays and the quasifocal line \mathbf{f} is shown in Figure 14 ($\alpha=60^\circ$). The parameters of form and location of the quasifocal line \mathbf{f} changed when the angle is changed. Incident rays are not shown in Figures 15.

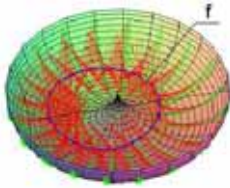


Figure 15: The quasifocal line \mathbf{f} with $\alpha=60^\circ$.

Parametric equation of the quasifocal line \mathbf{f} when the direction of incident rays is perpendicular to the axis of the torus (Fig.16) is

$$\begin{aligned} x &= x(u) = R[\sin(u), \\ z &= z(u) = R\cos(u) - \frac{r}{2\cos(u)}. \end{aligned} \quad (8)$$

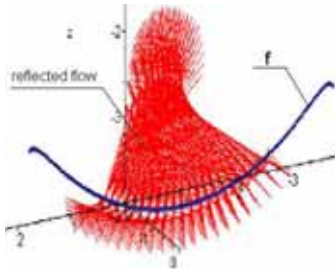


Figure 16: Reflected flow and the quasifocal line \mathbf{f} for the torus reflecting surface.

3. ENERGY MAP

On the physical models the temperature distribution gives us the possibility to determine hot spots which provoke stress in the receiver and can cause material failure.

The temperature on the receiver corresponds to the concentration ratio of the reflected rays flux.

The calculations based on the energy maps obtained with mathematical and computer simulation (Fig 16).

The reflecting properties analysis of surfaces was done on a base of the quasifocal points and lines. The geometrical model forming of a reflecting surface on a given quasifocal curve is considered. That presents the opportunities in designing of reflecting systems in view of required parameters of a concentration zone.

This proposed ways give the possibility to design reflecting systems with reflectors using large class of surfaces: rotation surfaces with demand generators; canal surfaces with demand focal zone; regular spiroid surfaces with a demand generators and so on.

To simulate a surface of a receiver with the most absorbing of power the geometrical apparatus is proposed.

4. CONCLUSIONS

The effectiveness of a concentrating system depends on form and location parameters of a receiver. The theory of the quasifocal lines gives us the possibility to determine zones of maximum energy concentration and to locate receivers with correspondent parameters.

The computer simulation gives us the possibility to obtain concentrating system exactly and visually and calculate the concentration ratios for different reflectors. It is important also to have equations of quasifocal lines for concentrating systems with stationary concentrators.

The flux distribution on the surface of a receiver is determined. The calculations are

based on the energy maps obtained with mathematical and computer simulation. Three dimensions model as the surface of flux distribution for quantitative estimation of flux distribution on the receiver is obtained. The lines of equal energy on a receiver for a torus reflecting surface were constructed (Fig.17). It is also important to know the secondary reflecting flow properties and flux distribution after the secondary reflection.

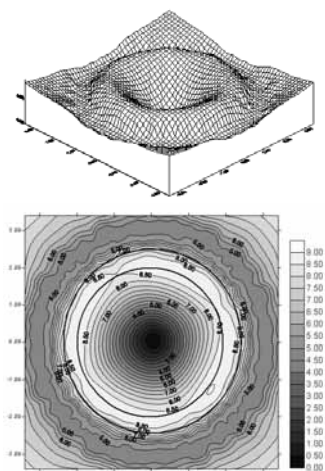


Figure 17: Isolines of local concentration ratio for the torus reflected surface.

ACKNOWLEDGMENTS

The authors express their thanks to the reviewers for their useful comments.

REFERENCES

- [1] Brocard H. and Lemoyne T. *Courbes Geometriques Remarquables (courbes speciales) planes et gauches*, Vol. 1. Paris: Blanchard. 1967.
- [2] Cayley A. A memoir upon caustics. *Phil. Trans. Roy. Soc. Lond.* 147:273-312; *Collected Works*. 1857, Vol. 2. pp. 336-380.
- [3] G. Glaeser: *Reflections on Spheres and Cylinders of Revolution*. *J. for Geometry and Graphics*. Volume 3, 121-139 (1999).

[4] Pidgorny O. L. *Constructive Modeling of Bundles Chains and Congruencies of Normals as Result of Reflection and Refraction*. The 10th international conference on geometry and graphics. Volume 1.- Kyiv, 2002.-p. 222-225.

[5] Bruce J.W., Giblin P.J., Gibson C.G. *On caustics by reflection*, *Topology*. 1981. 1.21. 179 – 199.

[6] Dvoretzky A.T. *Geometrical models of solar energy reflected flows*. The 10th international conference on geometry and graphics. Volume 1.- Kyiv, 2002.-p. 222-225.

[7] Dvoretzky A. T. *Solar Thermal Concentrating Technologies in Installations with Stationary Concentrators*. The XIIth SolarPACES International Symposium on Concentrating Solar Power and Chemical Energy Systems.-Seville, Spain.2006.

[8] Dvoretzky A.T. *Apparatus of Secondary Reflection*. The 13th International Conference on Geometry and Graphics. Dresden, Germany, (2008) 81- 85.

[9] Dvoretzky A.T. *Reflected Surfaces and Their Directing Cones*. *J. for Geometry and Graphics*. Volume14(2010), Number 1, 69-80.

[10] Denysova T.V. *Focal surface of reflected rays congruence for surfaces of revolution*. *Energy-savings in civil engineering and architecture*.- Kyiv: - 2011 – Issue No 2, 66-70.

ABOUT THE AUTHORS

1. Dvoretzky Alexander. The head of Geometrical and Computer Simulation Department, Academy of Environmental Protection and Resort Development, Crimea, Russian Federation. Doctor, professor. E-mail: dvoretzky@napks.edu.ua and erces_crimea@mail.ru

2. Denysova Tetiana. Assistant of professor of Geometrical and Computer Simulation Department, Academy of Environmental Protection and Resort Development, Crimea, Russian Federation. E-mail: smorodina82@mail.ru

CLASSIFICATION OF MANIFOLDS RESULTING AS MINKOWSKI OPERATION PRODUCTS OF BASIC GEOMETRIC POINT SETS

Daniela VELICHOVÁ

Slovak University of Technology in Bratislava, Slovakia

ABSTRACT: This paper brings basic classification of possible results of point set operations Minkowski sum and Minkowski product performed on elementary geometric figures as point subsets in the Euclidean space \mathbf{E}^n , which can be represented analytically by the relevant vector maps. Intrinsic geometric properties of resulting manifolds are presented with illustrations of various appealing results in \mathbf{E}^3 and 3-dimensional orthographic views of manifolds from spaces \mathbf{E}^4 and \mathbf{E}^6 .

Keywords: Minkowski point set operation, vector sum, exterior wedge product of vectors, Minkowski point set combination, Minkowski operator.

1. INTRODUCTION

In accordance with the original definition of Minkowski sum of two point sets based on their position with respect to the fixed reference point, both Minkowski point set operations are introduced as operations defined on vectors determined by their Cartesian coordinates. Minkowski sum of point sets is defined as vector sum and Minkowski product of point sets as exterior (wedge) product of position vectors of all points in the respective sets. Vector maps of operands in the two Minkowski point set operations determine form of the vector maps of the resulting manifolds. Therefore, they can be represented analytically and their intrinsic geometric properties can be derived by means of differential characteristics of the two operand manifolds. Different properties of background vector operations of vector sum and wedge product of vectors influence intrinsic characteristics of resulting manifolds, too.

Position of operand sets with respect to the reference point at the origin O of the Cartesian coordinate system is one of the key characteristics of the classifications. Let us suppose first that none of the operand manifolds determined by its vector map defined on interval in real numbers contains reference point O . Such situations will be considered as special singu-

larities and treated separately. Assuming Cartesian coordinates of an arbitrary point $p \neq O$ from \mathbf{E}^n , $p = (p_1, p_2, \dots, p_n)$, its non-zero position vector be $\mathbf{p} = (p_1, p_2, \dots, p_n)$. Manifolds represented by their vector maps are considered as infinite sets of points whose position vectors are determined as respective values of the vector functions defining the manifolds. Therefore, operations of Minkowski sum and Minkowski product of two manifolds can be performed as operations on their vector maps.

Starting from Minkowski sum and Minkowski product of two points, the resulting points are determined straightforwardly, while sum of points $a, b \in \mathbf{E}^n$ is again a point in \mathbf{E}^n , whereas their product is a point in the space \mathbf{E}^d , for $d = n(n-1)/2$

$$a = (a_1, a_2, \dots, a_n), b = (b_1, b_2, \dots, b_n)$$

$$a + b = (a_1 + b_1, a_2 + b_2, \dots, a_n + b_n)$$

$$a \wedge b = (a_1 \mathbf{e}_1 + \dots + a_n \mathbf{e}_n) \wedge (b_1 \mathbf{e}_1 + \dots + b_n \mathbf{e}_n) =$$

$$= (a_1 b_2 - a_2 b_1) \mathbf{e}_{1,2} + \dots + (a_{n-1} b_n - a_n b_{n-1}) \mathbf{e}_{n-1,n}.$$

Bivectors $\mathbf{e}_{1,2}, \mathbf{e}_{1,3}, \dots, \mathbf{e}_{1,n}, \mathbf{e}_{2,3}, \dots, \mathbf{e}_{2,n}, \dots, \mathbf{e}_{n-1,n}$ form ortho-normal basis of the space $\wedge^2(\mathbf{E}^3)$. For dimension $n = 3$, wedge product can be considered as equivalent to the vector product. Euclidean space \mathbf{E}^3 with basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ is isomorphic to the space $\wedge^2(\mathbf{E}^3)$ with

basis $\{\mathbf{e}_1 \wedge \mathbf{e}_2, \mathbf{e}_1 \wedge \mathbf{e}_3, \mathbf{e}_2 \wedge \mathbf{e}_3\}$, because there exists a regular linear transformation $\varphi: \wedge^2(\mathbf{E}^3) \rightarrow \mathbf{E}^3$ mapping one basis to the other one. Matrix of this transformation is

$$M_\varphi = \begin{pmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

For an arbitrary vector \mathbf{u} holds $\varphi(\mathbf{u}) = \mathbf{u} \cdot M_\varphi$, therefore $\varphi(\mathbf{e}_1 \wedge \mathbf{e}_2) = \mathbf{e}_3$, $\varphi(\mathbf{e}_1 \wedge \mathbf{e}_3) = -\mathbf{e}_2$, and $\varphi(\mathbf{e}_2 \wedge \mathbf{e}_3) = \mathbf{e}_1$.

For vector product of vectors \mathbf{u} and \mathbf{v} holds

$$\begin{aligned} \mathbf{u} \times \mathbf{v} &= (u_2 v_3 - u_3 v_2) \mathbf{e}_1 + (u_3 v_1 - u_1 v_3) \mathbf{e}_2 + \\ &\quad + (u_1 v_2 - u_2 v_1) \mathbf{e}_3 \\ \varphi(\mathbf{u} \wedge \mathbf{v}) &= \varphi \left(\begin{pmatrix} (u_1 v_2 - u_2 v_1)(\mathbf{e}_1 \wedge \mathbf{e}_2) + \\ (u_1 v_3 - u_3 v_1)(\mathbf{e}_1 \wedge \mathbf{e}_3) + \\ (u_2 v_3 - u_3 v_2)(\mathbf{e}_2 \wedge \mathbf{e}_3) \end{pmatrix} \right) = \\ &= (u_1 v_2 - u_2 v_1) \varphi(\mathbf{e}_1 \wedge \mathbf{e}_2) + \\ &\quad + (u_1 v_3 - u_3 v_1) \varphi(\mathbf{e}_1 \wedge \mathbf{e}_3) + \\ &\quad + (u_2 v_3 - u_3 v_2) \varphi(\mathbf{e}_2 \wedge \mathbf{e}_3) = \\ &= (u_1 v_2 - u_2 v_1) \mathbf{e}_3 - (u_1 v_3 - u_3 v_1) \mathbf{e}_2 + \\ &\quad + (u_2 v_3 - u_3 v_2) \mathbf{e}_1 = \mathbf{u} \times \mathbf{v}. \end{aligned}$$

Inverse linear transformation φ^{-1} to the transformation φ yields $\varphi^{-1}(\mathbf{u} \times \mathbf{v}) = \mathbf{u} \wedge \mathbf{v}$, while for its matrix holds $M_{\varphi^{-1}} = M_\varphi^{-1}$.

Wedge product and vector product of vectors in \mathbf{E}^3 are equivalent and Minkowski product of two points in \mathbf{E}^3 is therefore a point in \mathbf{E}^3 . Wedge product of two vectors in \mathbf{E}^4 is a vector in \mathbf{E}^6 , so Minkowski product of two points in \mathbf{E}^4 is a point in \mathbf{E}^6 .

Let a differentiable manifold M be determined in the space \mathbf{E}^n by its vector map

$$\mathbf{r}(u_i) = (x_1(u_i), x_2(u_i), \dots, x_n(u_i)), u_i \in \Omega,$$

and coordinate functions $x_1(u_i), x_2(u_i), \dots, x_n(u_i)$ are defined and at least once differentiable on the region $\Omega \subset \mathbf{R}^i, i \leq n$. Minkowski sum of arbitrary point $p = [p_1, p_2, \dots, p_n]$ and manifold M is a manifold $M' \subset \mathbf{E}^n$, which is a manifold M translated by position vector \mathbf{p} .

Minkowski product of manifold M and point p is manifold $M^* \subset \mathbf{E}^d$, determined by vector map in the form

$$p \wedge M =$$

$$\begin{aligned} (p_1 \mathbf{e}_1 + \dots + p_n \mathbf{e}_n) \wedge (x_1(u_i) \mathbf{e}_1 + \dots + x_n(u_i) \mathbf{e}_n) = \\ = (p_1 x_2(u_i) - p_2 x_1(u_i), \dots, p_{n-1} x_n(u_i) - p_n x_{n-1}(u_i)) \end{aligned}$$

Examples from \mathbf{E}^3 are presented in Figure 1, where Minkowski products of a point and curve, point and planar region and point and surface patch are presented.

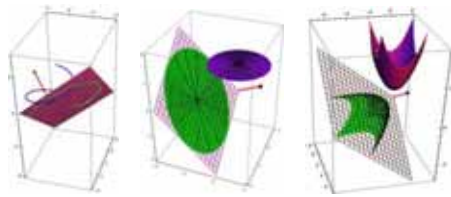


Figure 1: Minkowski product of point and manifolds – helix, disc, and parabolic patch.

Minkowski product of point p and manifold M is an image of M under a “quasi-central projection” from the centre p to the plane passing through origin O and perpendicular to the position vector \mathbf{p} of the centre p . Let $p = [p_1, p_2, p_3]$, and manifold M be determined by vector map

$$\mathbf{r}(u, v) = (x_1(u, v), x_2(u, v), x_3(u, v))$$

for $(u, v) \in \Omega \subset \mathbf{R}^2$. Minkowski product of M and p is a manifold $M^* = p \wedge M$ determined by

$$\mathbf{r}^*(u, v) = \mathbf{r}(u, v) \cdot \mathbf{T} = \mathbf{r}(u, v) \cdot \begin{pmatrix} 0 & p_3 & -p_2 \\ -p_3 & 0 & p_1 \\ p_2 & -p_1 & 0 \end{pmatrix}$$

where \mathbf{T} is the matrix representing this linear transformation. Image M^* is planar figure located in the plane $p_1 x + p_2 y + p_3 z = 0$.

Special positions of point p on coordinate axes yield different quasi-central projections to perpendicular coordinate planes that are linear transformations composed from orthographic projection to respective coordinate plane xy, xz, yz , revolution by angle $-\pi/2$ about origin O in this plane and scaling by nonzero coordinate of the centre of projection p_3, p_2, p_1 , with matrices

$$\begin{pmatrix} 0 & p_3 & 0 \\ -p_3 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & -p_2 \\ 0 & 0 & 0 \\ p_2 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & p_1 \\ 0 & -p_1 & 0 \end{pmatrix}.$$

In Figure 2 an image of hyperbolic paraboloid is presented under the quasi-central projection from point $p = [0, 0, 1]$ onto plane xy .

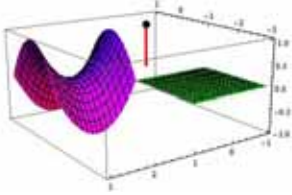


Figure 2: Minkowski product of point and surface patch.

Let a 2-dimensional manifold in E^4 be defined by map $\mathbf{r}(u, v) = (u, v, u^2v, u^2v^2)$, $(u, v) \in R^2$. 3D orthographic views of this surface patch in coordinate subspaces yzw , xzw , xyw and view under the quasi-central projection from the point $p = [0, 0, 0, 1]$ to the coordinate subspace xyz (representing revolved orthographic view to this coordinate space) are presented in Figure 3.

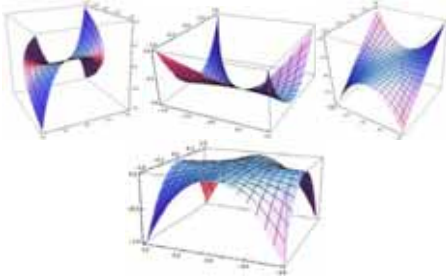


Figure 3: 3D views of a manifold from E^4 .

2. MINKOWSKI SUM AND PRODUCT OF LINES AND PLANES

2.1 Two lines

Consider now two lines in space E^3 given by parametric representations

$$\mathbf{p}(u) = p + u\mathbf{a}, \mathbf{q}(v) = q + v\mathbf{b}, u, v \in R.$$

Their position determines form of their Minkowski sum

$$S = \mathbf{p}(u) \oplus \mathbf{q}(v) = p + q + u\mathbf{a} + v\mathbf{b},$$

and Minkowski product for $(u, v) \in R^2$

$$P = \mathbf{p}(u) \otimes \mathbf{q}(v) = p \wedge q + p \wedge u\mathbf{a} + q \wedge v\mathbf{b} + u\mathbf{a} \wedge v\mathbf{b}.$$

Two parallel lines with collinear direction vectors $\mathbf{b} = k\mathbf{a}$ determine Minkowski sum as line S in the same direction passing through the point $p + q$. In case of one line passing through origin, Minkowski sum coincides with the other line.

Singular forms of Minkowski product are the following. If lines form a plane passing through origin, their Minkowski product is line passing through origin perpendicularly to this plane, i.e. to both lines. Minkowski product of line with itself is perpendicular line passing through origin if line does not contain origin, and it is just one point - the origin, if it does.

Suppose that lines do not form plane passing through origin. Their Minkowski product is plane passing through the point $p \wedge q$ and perpendicular to both lines, with the general equation

$$a_1x + a_2y + a_3z + a_3(q_1p_2 - q_2p_1) + a_2(q_3p_1 - q_1p_3) + a_1(q_2p_1 - q_1p_2) = 0$$

where

$$\mathbf{a} = (a_1, a_2, a_3), p = [p_1, p_2, p_3], q = [q_1, q_2, q_3].$$

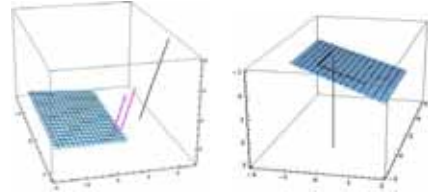


Figure 4: Minkowski sum and product of parallel line segments and line segments intersecting at origin.

Minkowski sum of intersecting lines with non-collinear direction vectors is plane S passing through the point $p + q$ and parallel to the plane formed by these lines. In case they form plane passing through origin, their Minkowski sum is the plane they form, see Figure 4, right.

Let lines do not form plane passing through origin. Then they form the plane determined by their common point p and their direction vectors \mathbf{a} and $\mathbf{b} = (b_1, b_2, b_3)$, with equation

$$(a_2b_3 - a_3b_2)x + (a_3b_1 - a_1b_3)y + (a_1b_2 - a_2b_1)z - (a_2b_3 - a_3b_2)p_1 - (a_3b_1 - a_1b_3)p_2 - (a_1b_2 - a_2b_1)p_3 = 0,$$

from which follows the condition satisfied for plane not passing through origin

$$(a_2b_3 - a_3b_2)p_1 + (a_3b_1 - a_1b_3)p_2 + (a_1b_2 - a_2b_1)p_3 \neq 0.$$

Minkowski product of such two lines is a quadratic surface with the equation

$$\begin{aligned} & a_1b_1x^2 + a_2b_2y^2 + a_3b_3z^2 + (a_1b_2 + b_1a_2)xy + \\ & + (a_3b_1 + a_1b_3)xz + (a_2b_3 + a_3b_2)yz + \\ & + \left[p_1^2(b_2a_3 - a_2b_3) + \right. \\ & \left. + p_1p_2(a_1b_3 - b_1a_3) + p_1p_3(b_1a_2 - a_1b_2) \right] x + \\ & + \left[p_2^2(a_1b_3 - b_1a_3) + \right. \\ & \left. + p_1p_2(b_2a_3 - a_2b_3) + p_2p_3(b_1a_2 - a_1b_2) \right] y + \\ & + \left[p_3^2(b_1a_2 - a_1b_2) + \right. \\ & \left. + p_1p_3(b_2a_3 - a_2b_3) + p_2p_3(a_1b_3 - b_1a_3) \right] z = 0 \end{aligned}$$

Using standard methods it can be proved that it is general equation of a hyperbolic paraboloid.

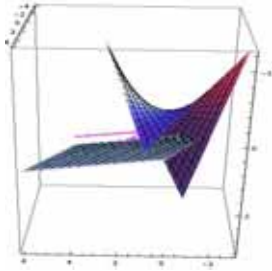


Figure 5: Minkowski sum and product of intersecting line segments.

In singular situation, when two intersecting lines form a plane passing through origin, their Minkowski product is a line passing through origin and perpendicular to the plane formed by the two lines, see in Figure 4, right.

Minkowski sum of two skew lines with non-collinear direction vectors \mathbf{a} , \mathbf{b} passing through points p and q is plane S determined by the two vectors passing through the point $p + q$. If one of the lines is passing through origin, say $p = O$, then Minkowski sum

$S = \mathbf{p}(u) \oplus \mathbf{q}(v) = q + u\mathbf{a} + v\mathbf{b}$, $u, v \in \mathbf{R}$ is plane passing through the other line, see in Figure 6, on the right.

Minkowski product of two skew lines is a plane if and only if one of the lines is parallel to the plane passing through origin and the other line. In case one from the skew lines is passing through origin, $p = O$, their Minkowski product

$P = \mathbf{p}(u) \otimes \mathbf{q}(v) = q \wedge v\mathbf{b} + u\mathbf{a} \wedge v\mathbf{b} = (q + u\mathbf{a}) \wedge v\mathbf{b}$, $(u, v) \in \mathbf{R}^2$ is plane through origin perpendicular to this line, see in Figure 6, on the right. In all other configurations it is a hyperbolic paraboloid.

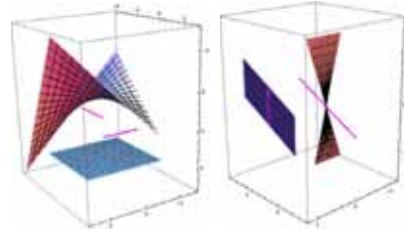


Figure 6: Minkowski sum and product of skew line segments.

2.2 Line and plane

Let line and plane be defined parametrically as

$$\mathbf{p}(t) = p + t\mathbf{a}, t \in \mathbf{R},$$

$$\mathbf{q}(u, v) = q + u\mathbf{b} + v\mathbf{c}, \quad u, v \in \mathbf{R}.$$

In case they are parallel, line direction vector can be represented as linear combination of plane direction vectors, $\mathbf{a} = k\mathbf{b} + l\mathbf{c}$, for $k, l \in \mathbf{R}$, and their Minkowski sum for $t, u, v \in \mathbf{R}$

$S = \mathbf{p}(t) \oplus \mathbf{q}(u, v) = p + q + (tk + u)\mathbf{b} + (tl + v)\mathbf{c}$ is plane parallel to both line and plane passing through the point $p + q$, provided none of them contains origin. If line (plane) contains origin, their Minkowski sum is the same plane (parallel plane passing through line), Figure 7.

Minkowski product of line and plane,

$$P = \mathbf{p}(t) \otimes \mathbf{q}(u, v) = \\ = p \wedge q + p \wedge (u\mathbf{b} + v\mathbf{c}) - q \wedge t(k\mathbf{b} + l\mathbf{c})$$

is a hyperbolic paraboloid, if plane is not passing through origin. In such case, Minkowski product of line parallel to plane is plane perpendicular to the line, see Figure 7, right bottom.

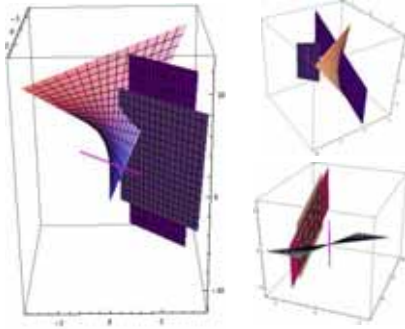


Figure 7: Minkowski sum and product of line and parallel plane.

Line and not parallel plane determine Minkowski sum

$S = \mathbf{p}(t) \oplus \mathbf{q}(u, v) = p + q + t\mathbf{a} + u\mathbf{b} + v\mathbf{c}$ forming the whole space \mathbf{E}^3 for $t, u, v \in \mathbf{R}$.

Their Minkowski product

$$P = \mathbf{p}(t) \otimes \mathbf{q}(u, v) = \\ = p \wedge q + (p + t\mathbf{a}) \wedge (u\mathbf{b} + v\mathbf{c}) + t\mathbf{a} \wedge q$$

is the same space \mathbf{E}^3 , if line is not passing through the origin. If just plane is passing through origin, Minkowski product is formed as 1-parametric system of planes with common line passing through origin and perpendicular to the line, see in Figure 8, on the top. If just line is passing through origin, then their Minkowski product is plane passing through origin and perpendicular to this line

$P = t\mathbf{a} \wedge (q + u\mathbf{b} + v\mathbf{c})$, $t, u, v \in \mathbf{R}$, see in Figure 8, bottom left. If line and plane share origin as their common point, the resulting plane intersects given plane in line passing through origin and perpendicular to the given line, Figure 8, bottom right.

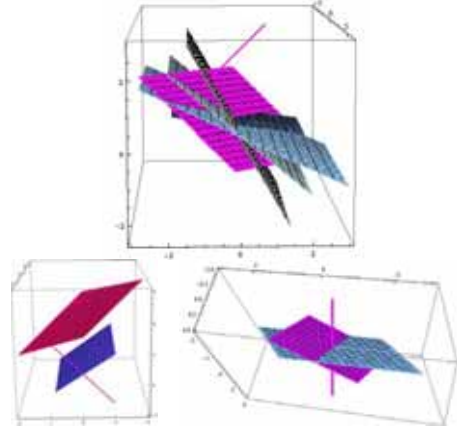


Figure 8: Minkowski product of plane and intersecting line.

2.3 Two planes

Let two intersecting planes in \mathbf{E}^3 be determined parametrically by vector maps

$$\mathbf{p}(u, v) = p + u\mathbf{a} + v\mathbf{b}, \quad u, v \in \mathbf{R},$$

$$\mathbf{q}(s, t) = q + s\mathbf{a} + t\mathbf{c}, \quad s, t \in \mathbf{R}.$$

Their Minkowski sum

$$S = \mathbf{p}(u, v) \oplus \mathbf{q}(s, t) = \\ = p + q + (u + s)\mathbf{a} + v\mathbf{b} + t\mathbf{c}$$

and Minkowski product

$$P = \mathbf{p}(u, v) \otimes \mathbf{q}(s, t) = \\ = p \wedge q + p \wedge (s\mathbf{a} + t\mathbf{c}) - \\ - (q \wedge (u\mathbf{a} + v\mathbf{b})) + ut(\mathbf{a} \wedge \mathbf{c}) + \\ + sv(\mathbf{a} \wedge \mathbf{b}) + vt(\mathbf{b} \wedge \mathbf{c})$$

form the whole space \mathbf{E}^3 . In case, both planes are passing through origin, their Minkowski product is one parametric system of hyperbolic paraboloids with common line in the pierce line of the two planes, passing through origin, see Figure 9, top left.

If the planes are parallel,

$$\mathbf{p}(u, v) = p + u\mathbf{a} + v\mathbf{b}, \quad u, v \in \mathbf{R},$$

$$\mathbf{q}(s, t) = q + s\mathbf{a} + t\mathbf{b}, \quad s, t \in \mathbf{R},$$

their Minkowski sum

$$S = \mathbf{p}(u, v) \oplus \mathbf{q}(s, t) = \\ = p + q + (u + s)\mathbf{a} + (v + t)\mathbf{b}$$

is a plane passing through the point $p + q$ in direction of both parallel planes.

Minkowski product of two parallel planes is the whole space \mathbf{E}^3 in case none of the planes is passing through origin,

$$P = \mathbf{p}(u, v) \otimes \mathbf{q}(s, t) = p \wedge q + p \wedge (sa + tb) - (q \wedge (ua + vb)) + (ut - vs)(a \wedge b),$$

formed by system of planes, see Figure 9, bottom. If one from the parallel planes is passing through origin, space is formed by system of lines in direction perpendicular to both planes, see Figure 9, top right.

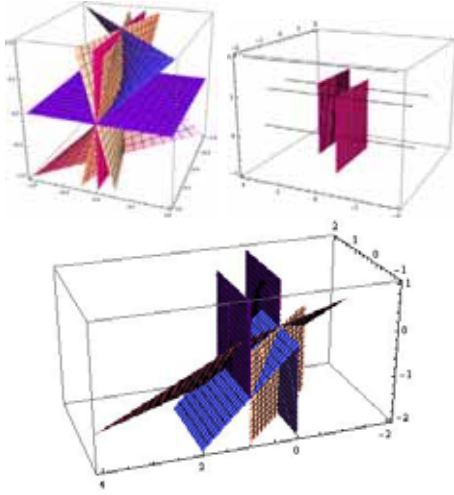


Figure 9: Minkowski product of two planes.

3. MINKOWSKI SUM AND PRODUCT OF CURVES IN 3D

Consider two curve segments determined parametrically

$$\mathbf{k}(u) = (xk(u), yk(u), zk(u)), u \in I \subset \mathbf{R}$$

$$\mathbf{l}(v) = (xl(v), yl(v), zl(v)), v \in L \subset \mathbf{R}.$$

Minkowski sum of the two curves is a translation surface patch defined on planar region $\Omega = I \times L \subset \mathbf{R}^2$

$\mathbf{s}(u, v) = (xk(u) + xl(v), yk(u) + yl(v), zk(u) + zl(v))$ that can be generated by translation of one curve along the other one. Differential characteristics of such surface patch can be represented by means of derivatives of the curve vector maps.

$$\mathbf{k}' = \mathbf{k}'(u) = (xk'(u), yk'(u), zk'(u))$$

$$\mathbf{l}' = \mathbf{l}'(v) = (xl'(v), yl'(v), zl'(v)).$$

$$\varphi_1 = \|\mathbf{k}'\|^2 du^2 + 2\mathbf{k}' \cdot \mathbf{l}' du dv + \|\mathbf{l}'\|^2 dv^2$$

$$D_1 = (\|\mathbf{k}'\| \cdot \|\mathbf{l}'\|)^2 - (\mathbf{k}' \cdot \mathbf{l}')^2 = \|\mathbf{k}' \times \mathbf{l}'\|^2$$

$$\varphi_2 = L du^2 + N dv^2, L = \frac{[\mathbf{k}', \mathbf{l}', \mathbf{k}']}{\|\mathbf{k}' \times \mathbf{l}'\|}, N = \frac{[\mathbf{k}', \mathbf{l}', \mathbf{l}']}{\|\mathbf{k}' \times \mathbf{l}'\|}$$

$$D_2 = L \cdot N = \frac{D}{D_1}, D = \begin{vmatrix} \|\mathbf{k}'\|^2 & \mathbf{k}' \cdot \mathbf{l}' & \mathbf{k}' \cdot \mathbf{l}'' \\ \mathbf{k}' \cdot \mathbf{l}' & \|\mathbf{l}'\|^2 & \mathbf{l}' \cdot \mathbf{l}'' \\ \mathbf{k}' \cdot \mathbf{l}'' & \mathbf{k}'' \cdot \mathbf{l}' & \mathbf{k}'' \cdot \mathbf{l}'' \end{vmatrix}$$

Gauss curvature is given by formula $K = \frac{D}{D_1^2}$.

Examples of Minkowski sum of line segment and circle are presented in Figure 10. Planar regions presented on the left and in the middle can be generated as Minkowski sum of a line segment parallel to (located in) the plane of the circle, elliptical cylindrical surface patch for non-parallel line segment.



Figure 10: Minkowski sum of circle and line segment in various positions.

Minkowski product of the two curves is a translation surface patch defined on planar region $\Omega = I \times L \subset \mathbf{R}^2$

$$\mathbf{p}(u, v) = \mathbf{k}(u) \otimes \mathbf{l}(v), (u, v) \in \Omega$$

$$\mathbf{p}(u, v) = \begin{pmatrix} yk(u)zl(v) - zk(u)yl(v) \\ zk(u)xl(v) - xk(u)zl(v) \\ xk(u)yl(v) - yk(u)xl(v) \end{pmatrix}^T$$

generated by all such points in the space, whose position vector is vector product of position vectors of one point from curve \mathbf{k} and one point from curve \mathbf{l} . In case of elementary planar curves well-known surface patches are generated, as e.g. ruled surfaces - cylinders, transition surfaces, conoids, looped strips, torus, etc.

Using properties of vector product, Lagrange identity and following abbreviations

$\mathbf{k}' \times \mathbf{l} = \mathbf{a}, \mathbf{k} \times \mathbf{l}' = \mathbf{b}, \mathbf{k}' \times \mathbf{l}' = \mathbf{c}, \mathbf{k}'' \times \mathbf{l} = \mathbf{d}, \mathbf{k} \times \mathbf{l}'' = \mathbf{e}$
differential characteristics of Minkowski product of two curves can be derived in forms

$$\varphi_1 = \|\mathbf{a}\|^2 du^2 + 2\mathbf{a} \cdot \mathbf{b} du dv + \|\mathbf{b}\|^2 dv^2$$

$$D_1 = \|\mathbf{a} \times \mathbf{b}\|^2$$

$$\varphi_2 = D_1^{-1} (L du^2 + 2M du dv + N dv^2)$$

$$L = [\mathbf{d}, \mathbf{a}, \mathbf{b}], M = [\mathbf{c}, \mathbf{a}, \mathbf{b}], N = [\mathbf{e}, \mathbf{a}, \mathbf{b}]$$

$$D_2 = D_1^{-1} (D_{LN} - D_M)$$

$$D_{LN} = \begin{vmatrix} \mathbf{d} \cdot \mathbf{e} & \mathbf{d} \cdot \mathbf{a} & \mathbf{d} \cdot \mathbf{b} \\ \mathbf{a} \cdot \mathbf{e} & \|\mathbf{a}\|^2 & \mathbf{a} \cdot \mathbf{b} \\ \mathbf{b} \cdot \mathbf{e} & \mathbf{a} \cdot \mathbf{b} & \|\mathbf{b}\|^2 \end{vmatrix}, D_M = \begin{vmatrix} \|\mathbf{c}\|^2 & \mathbf{c} \cdot \mathbf{a} & \mathbf{c} \cdot \mathbf{b} \\ \mathbf{a} \cdot \mathbf{c} & \|\mathbf{a}\|^2 & \mathbf{a} \cdot \mathbf{b} \\ \mathbf{b} \cdot \mathbf{c} & \mathbf{a} \cdot \mathbf{b} & \|\mathbf{b}\|^2 \end{vmatrix}$$

Gauss curvature is given by formula

$$K = \frac{D_{LN} - D_M}{D_1^2}$$

Examples of Minkowski products of a line segment and a circle that are located in different super-positions are patches illustrated in Figure 11. For the patch on the left, the line segment is perpendicular to the osculating plane of the circle, line segment and circle are located in the same plane for the transition patch in the middle, and 2-sided strip on the right is the Minkowski product of a line segment parallel to the circle osculating plane.



Figure 11: Minkowski products of circle and line segment in various superpositions.

Other interesting examples are Minkowski sum and product of two circles. Positioning circles into different planes, surface patches of interesting aesthetic forms can be generated. Minkowski sum (disc) and product (torus) of circles in parallel planes are presented in Figure 12. Position of concentric circles in perpendicular planes results in surface patches presented

in Figure 13, Minkowski sum is on the left and Minkowski product is on the right. Different form of Minkowski product can be achieved with non-concentric circles, as illustrated in Figure 14, right, while their Minkowski sum on the left is identical, up to position in the space.

Interesting questions can arise in connection to identifying such points on resulting manifolds, whose curvilinear coordinates u and v are equal. These point sets can be represented as so called partial Minkowski operation results. Assuming both operand curves as parameterized for the same parameter and on the same interval $I \subset \mathbf{R}$, the resulting manifold is a curve, represented by its vector map on I

$$\mathbf{s}(u) = (xk(u) + xl(u), yk(u) + yl(u), zk(u) + zl(u))$$

$$\mathbf{p}(u) = \begin{pmatrix} yk(u)zl(u) - zk(u)yl(u) \\ zk(u)xl(u) - xk(u)zl(u) \\ xk(u)yl(u) - yk(u)xl(u) \end{pmatrix}^T$$

and located on the respective Minkowski sum or product of differently parameterized curves. Illustrations are in Figures 12 – 14 (in red), in some configurations these are sets of special points on surfaces, as outline of a planar region, neck circle on torus, set of parabolic points on surface. Their properties are yet not clearly determined and they are topic of current study.

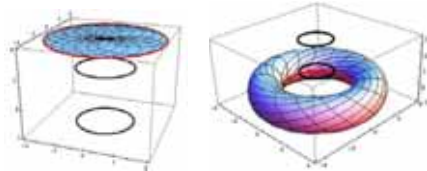


Figure 12: Circles in parallel planes.

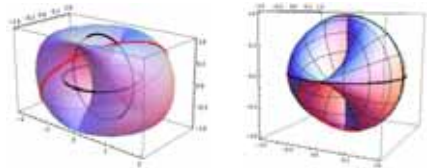


Figure 13: Circles in perpendicular planes I.

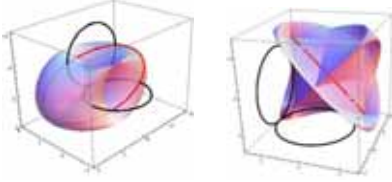


Figure 14: Circles in perpendicular planes II.

Illustration of curves generated as Minkowski product of two circles sharing none, one and two common points is in Figure 15. The resulting curve seems to have double (triple) point in the point corresponding to common points of operand curves.



Figure 15: Minkowski product of 2 equally parameterized circles in perpendicular planes.

Minkowski sum of two circles is an ellipse or a circle (for circles in parallel planes). Both curves are presented on respective surface patches in Figure 16.

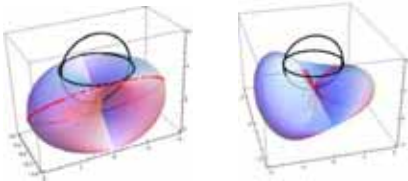


Figure 16: Minkowski sum and product of 2 circles in perpendicular planes.

4. MINKOWSKI SUM AND PRODUCT OF CURVES IN HIGHER DIMENSIONS

Minkowski sum of two curve segments in the space E^4 determined parametrically

$$\mathbf{k}(u) = (xk_1(u), xk_2(u), xk_3(u), xk_4(u)), u \in I$$

$\mathbf{l}(v) = (xl_1(u), xl_2(u), xl_3(u), xl_4(u)), u \in L$ can be considered as surface patch in E^4 with

parametric vector map

$$\mathbf{s}(u, v) = \mathbf{k}(u) \oplus \mathbf{l}(v) = \begin{pmatrix} xk_1(u) + xl_1(v) \\ xk_2(u) + xl_2(v) \\ xk_3(u) + xl_3(v) \\ xk_4(u) + xl_4(v) \end{pmatrix}$$

defined on $\Omega = I \times L$.

Orthographic views of resulting manifolds in 4 possible coordinate subspaces in E^4 provide 4 different forms of surface patches in 3D, as Minkowski sum of ellipse and conical helix illustrated in Figure 17.

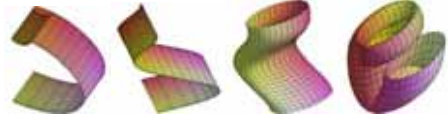


Figure 17: Minkowski sum of 2 curves in E^4 .

Minkowski product of curves $\mathbf{k}(u)$ and $\mathbf{l}(v)$ from E^4 generates a 2-parametric manifold in the space E^6 with vector map

$$\mathbf{p}(u, v) = \mathbf{k}(u) \otimes \mathbf{l}(v) = \begin{pmatrix} xk_1(u)xl_2(v) - xk_2(u)xl_1(v) \\ xk_1(u)xl_3(v) - xk_3(u)xl_1(v) \\ xk_1(u)xl_4(v) - xk_4(u)xl_1(v) \\ xk_2(u)xl_3(v) - xk_3(u)xl_2(v) \\ xk_2(u)xl_4(v) - xk_4(u)xl_2(v) \\ xk_3(u)xl_4(v) - xk_4(u)xl_3(v) \end{pmatrix}$$

defined on $\Omega = I \times L$.

Differential characteristics of both manifolds in higher dimensions and their dependence on the characteristics of operand curve segments are interesting topics for further theoretical study. Surface patches in E^6 generated as Minkowski product of two curves in E^4 can be successively orthographically mapped into various 5, 4 and 3-dimensional subspaces, leading to a variety of new surface forms in E^3 .

Two circles located in different coordinate planes in E^4 taken as operands of Minkowski product determine two-dimensional manifold in E^6 . Some of its orthographic views in different three-dimensional coordinate subspaces are illustrated in Figure 18.

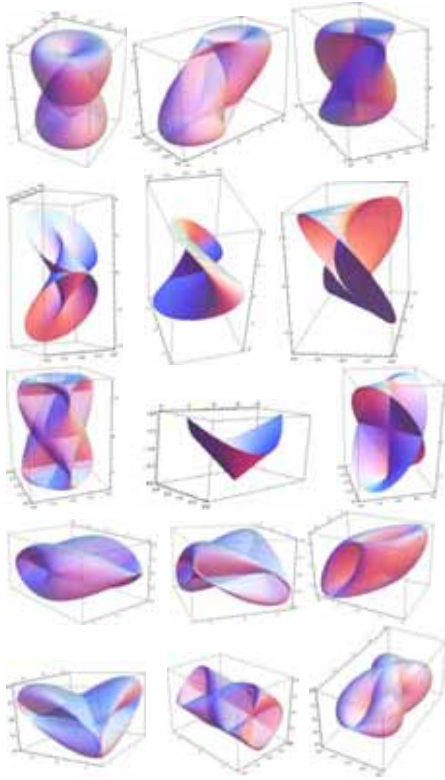


Figure 18: Different orthographic views in \mathbf{E}^3 of one surface patch from \mathbf{E}^6 .

Shamrock curve
 $\mathbf{r}(u) = (a_1 \cos u \sin^2 u, b_1 \sin u \cos^2 u, 0, 0), u \in \langle 0, 2\pi \rangle$
 and versière
 $\mathbf{s}(v) = \left(0, a_2(2v-1), \frac{b_2}{c_2 + d_2(2v-1)^2}, 0 \right), v \in \langle 0, 1 \rangle$
 form Minkowski sum

$$\mathbf{s}(u, v) = \begin{pmatrix} a_1 \cos u \sin^2 u \\ b_1 \sin u \cos^2 u + a_2(2v-1) \\ \frac{b_2}{c_2 + d_2(2v-1)^2} \\ 0 \end{pmatrix}^T$$

and Minkowski product

$$\mathbf{p}(u, v) = \begin{pmatrix} a_1 a_2 (2v-1) \cos u \sin^2 u \\ a_1 \frac{b_2}{c_2 + d_2(2v-1)^2} \cos u \sin^2 u \\ 0 \\ b_1 \frac{b_2}{c_2 + d_2(2v-1)^2} \sin u \cos^2 u \\ 0 \\ 0 \end{pmatrix}^T$$

that are surface patches located in different 3-dimensional subspaces of \mathbf{E}^4 and \mathbf{E}^6 visualized in Figure 19.

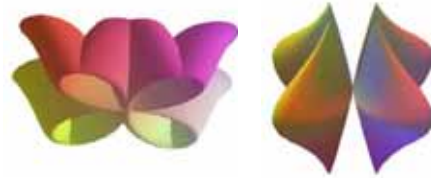


Figure 19: Minkowski sum and product.

5. CONCLUSIONS

Classification of possible results of Minkowski sum and product of basic geometric sets like points, lines and planes are presented in the paper, with illustrations of possible resulting manifolds in dependence on the position of the two operand sets. Basic differential characteristics are derived for manifolds resulting as Minkowski sum and product of two curves in the space \mathbf{E}^3 . Illustrations of Minkowski sum and product of two curves positioned into different planes in \mathbf{E}^3 and \mathbf{E}^4 are included as inspiration ideas. Presented set operations might be applied in design or architecture for modelling unusual shapes, in computer art for smooth modelling and morphing of aesthetic objects, and in various other areas of computer graphics and applications in art and science.

Concepts of Minkowski sum and Minkowski product of two point sets A and B can be also generalized to the concept of Minkowski set operators determining Minkowski combinations of point sets

$$LS_{k,l}(A, B) = kA \oplus lB, PS_{k,l}(A, B) = kA \otimes lB.$$

Thus a more powerful modelling tool is provided, in which k and l scalar multiples of point sets A and B are summed or multiplied and generate shapes resembling natural forms that might be applicable for purposes of bionics.

Images of Minkowski linear sum operator applied to two cylindrical helices for various coefficients are presented in Figure 20, on top, while illustration of their Minkowski product operator images are in Figure 20, bottom.

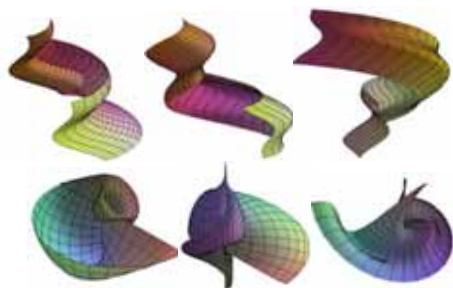


Figure 20: Minkowski sum and product combinations of 2 helices.

ACKNOWLEDGMENTS

This paper was supported by the project APVV-0161-12 grant awarded by Slovak research and development agency.

REFERENCES

- [1] J. Dobráková and V. Záhonová. Minkowski Product of Lines. In *G – Slovak Journal for Geometry and Graphics*, vol. 9, No. 18, pages 5-12, SSGG, 2012.
- [2] J. Dobráková and V. Záhonová. Minkowski Product of Point and Curve or Point and Surface in Three-Dimensional Euclidean Space. In *G - Slovak Journal for Geometry and Graphics*, vol. 11, No. 21, pages 17-26, SSGG, 2014.
- [3] A. Kaul and R. T. Farouki. Computing Minkowski Sums of Planar Curves. In *International Journal of Computational Geometry and Applications* 5, 1995, pages 413-432.
- [4] I. K. Lee and M. S. Kim and G. Elberg. The Minkowski Sum of 2D Curved Objects. In *Proceedings of Israel-Korea Bi-National Conference on New Themes in Computerized Geometrical Modelling*. Tel-Aviv Univ., pages 155-164, 1998.
- [5] D. Velichová. Multidimensional Riemannian Manifolds as Minkowski Products. In *Aplimat - Journal of Applied Mathematics*, N°2/2012, Volume 4, pages 151-158, Slovak University of Technology in Bratislava, 2012.
- [6] D. Velichová. Minkowski Set Operators. In *G - Slovak Journal for Geometry and Graphics*, vol. 9, No. 17, pages 53-66, SSGG, 2012.
- [7] D. Velichová. Minkowski Set Operations in Modelling of Manifolds. In *Proceedings of the GeoGra 2012 International Conference*. Budapest 2012, Hungary. CD-rom, 4 pages.
- [8] D. Velichová. The Art of Geometry. In *Proceedings 2013, 16th Conference Bridges Enschede*. Tesselations, Phoenix, Arizona, USA, pp. 143-150.
- [9] D. Velichová. Minkowski Operators in Shape Modelling. In *Mathematics and Art III, Proceedings of ESMA Conference*. Cagliari 2013, 8 pages. Available on-line at: math-art.eu/Cagliari2013-Lectures.php

ABOUT THE AUTHOR

Daniela Velichová is professor of applied mathematics at the Slovak University of Technology in Bratislava. Her professional interests cover applied mathematics, geometric modelling and differential geometry of curves and surfaces, projection methods in spaces of higher dimension, didactics of mathematics and geometry, and connections between mathematics and art. She is involved in several scientific projects investigating Minkowski point set operations and projects on Maths and Art.

COEXISTENCE OF GEOMETRY AND GEOLOGY - RECONSIDERING LE CORBUSIER

Michio KATO

The University of Tokyo, Japan

ABSTRACT: Le Corbusier, who is known as a modern architect, was also a painter and started his career of a painter as Purist, summarized his activity as a painter late in life as follows. “I have drawn or painted only women, or images, or symbols, or geology of women”. However he had exclusively painted the industrial commodity or musical instruments called <object-type>. How should we think his paintings of Purism? In order to reply to this question, we proposed to oppose <ge-o-logy> with <ge-o-metry>. Based on this opposition, a series of his activity as a painter were summarized as follows. 1) <Female images> had existed from the beginning (before Purism). 2) <Female images> disappeared in the time of Purism. However, <female images> had survived behind them, transferring to daily necessities, such as glasses or musical instruments. 3) <Female images> returned after his independence from Purism. Opposition of <ge-o-logy> with <ge-o-metry> does not cover only his pictures. It is connected to his city design too. It corresponded to the opposition between the curves and the straight lines in his urban design principle, which was shown in his article “the pack-donkey’s way and man’s way”. Thus, if we looked back upon his total activity as painter, we could not consider that his re-introduction of <female images> was a simple denial of Purism. We had rather to consider that the world of coexistence, where <female images> intervened between <geology> and <geometry>, was created. Finally <geology> included <geometry>. The former corresponded to <female images> and the latter underlay the basis of suppression of Purism. The horizon let him say as follows. “I have drawn or painted only women or the images, or the symbols, or the geology of women”.

Keywords: Le Corbusier, geometry, geology

1. INTRODUCTION

Le Corbusier, who is known as a modern architect, was also a painter. He started his career of a painter as Purist. Purism was born in late 1918 as an inheritor of Cubism criticizing Cubism. Purists painted exclusively the industrial commodity or musical instruments called <object-type>, following geometrical rules. In other words Purist inhibited <female images>.

However he summarized his activity as a painter late in life as follows.

I have drawn or painted only women, or the images, or the symbols, or the geology of women [1].

How should we think his Purist paintings? In order to reply to this question, we proposed to oppose <ge-o-logy> against <ge-o-metry>. Because they have a same root ‘ge’ but different components, i.e. the former literally means “logic of the earth” and the latter means “metric (measurement or measuring method) of the earth”.

First we would like to pay our attention to his beginning and prohibition of <female images>. Le Corbusier at first made many genres of works in addition to still life which was the exclusive theme of Purism. His works covered portraits, landscapes, and religion (history). Especially, he repeated <female images> including female nudes. However, human images

disappeared in his works for a while after the establishment of Purism. And <female images> returned after his independence from Purism. And he changed his signature of pictures from “Jeanneret” to “Le Corbusier” in 1928.

Next, we would like to pay our attention to his transition toward mythological images.

On and after the 1930s, deformation of objects became grater in his pictures, including not only <female images> but <object types> (daily necessities). And his images were related with various different genres of objects. They were partly converted into mythological images during World War II. The mythological images became one of main postwar picture motifs.

Incidentally his <female images> and mythological images had duality in relation to real objects actually referred to. On the one hand he shortened the distance from the body of real (particular and specific) women which his picture's motif. On the other hand he kept his distance from the real objects, with deforming their original outlines.

In this research, we would like to reconsider <female images> in his works and understand the process of the duality.

2. BEGINNING AND PROHIBITION OF <FEMALE IMAGES>

2.1 Beginning of <female images>

In his 100 years anniversary, the exhibition “Le Corbusier pictures before purism” was held in La Chaux de Fonds where he was born. His works, shown in the exhibition catalog [2], demonstrate that Jeanneret (original his family name of “Le Corbusier” that is his pen name) made many female works, before he started his activity in Paris (Purism).

Many pencil study sketches of female nude were drawn in 1908. From 1912 to 3, in addition to “the women of the seashore” (1912), many watercolor works were made. Especially in 1917, just before advocating Purism with Ozenfant, he made many female nude works. The exhibition catalog shows 32 sketches

among 66 ones in his sketchbook “No 10” (1917). We can check at least six watercolor nude works (including gouaches) in 1917. It shows that female nude was repeated as his motif before Purism.

2.2 Prohibition of <female images>

In his Purist era he had not selected female nudes as his picture subject. The exclusion was thought to be the constraint on Purist rules made with Ozenfant. However in 1918 when they declared Purism in their book: *After cubism*, the constraint was not so severe. They allowed three genres of painting and ordered as follows; 1) figure, 2) still life, 3) landscape. There were portraits by them in their first exhibition held at Tomas gallery in 1918.

It was after the first exhibition that they became to limit their subjects to <object type> (industrial commodity or musical instruments). And the second Purist exhibition was held in Druet gallery in 1921. The exhibition was composed of their works after the first exhibition. There were shown the process of making rules of Purism. And the rules was written in “purism” in *L'esprit nouveau* (vol. 4, Jan., 1921) [3].

The above facts showed that he prohibited female nudes after he had started his career as Purist with Ozenfant. In other words the establishment of Purism with Ozenfant suppressed female nudes in the form of the constraint of picture subjects.

3. RETURN TO <FEMALE IMAGES>

3.1 Coexistence and juxtaposition of male properties and female properties

If we reconsider his return to <female images>, we cannot but notice that it is unexplainable as a simple renewal of the preceding the Purist rules.

The reason is as follows; 1) <female images>, especially those of female nude, adjoined sexual desire or mythological motifs, 2) they conflicted with the aesthetics of Purism, 3) they were independent of Purist rules.

So we should reconsider Le Corbusier's female motif, especially female nude, carefully.

First we should point out that <female images> were generally based on the property against the male.

The female properties and the male properties have been repeatedly mentioned in the field of architecture since Vitruvius of the Roman Empire. First they were related to the order of columns as Vitruvius pointed out. Secondly they became to be related to geometrical shapes. Straight lines and a right angle were connected to the male, and curves were connected to the female. Western architectural view has been based on the opposition of the male properties with the female properties. In that case, the former has been usually placed higher than the latter. This view corresponds to his first criticism on city planning, "the pack-donkey's way and man's way" in *L'esprit nouveau* (vol. 17, June, 1922) [4].

His hierarchical classification was seen in his pictures too. Geometry, which was behind Purism, could be related to the superiority of the male. His exclusion of <female images> in Purism was on the extension of this idea. However the graceful curves as those of a violin or a pitcher coexisted with geometric straight lines and a right angle in Purist paintings. And they had been the main components of Purist paintings.

Finally we could say as follows. His juxtaposition or coexistence of male properties and female properties were the basic nature of his pictures. This view is seen in his theory of urban design. The straight lines of "man's way" correspond to the male and the curves of "the pack-donkey's way" correspond to the female.

However, we should reconfirm as follows; those curves differ from the female curves. And they were not opposed to the priority of reason or geometry in Purism.

3.2 Autonomous female image

We should add another origin about his return to the <female images> in order to understand his <female images>. Another origin was

autonomous, which were brewed through his encounter with many particular women. They were real but independent from any males. We should look back upon from this viewpoint.

If we mention any particular woman related to his formation of <female images>, it would be Yvonne Gallis (1892-1957) first. She was a model from Côte d'Azur and has a glamorous body. It was around 1921 when he met her for the first time. It was before he founded an architectural design office with his cousin, Pierre Jeanneret, in Paris. And she became his wife in 1930. Their relation continued till she died in 1957. There were many works of Yvonne.

The next person whom we should mention is Charlotte Perriand (1903-1998). She was employed as an assistant in October of 1927, and mainly took charge of the furniture design. Her collaboration with him continued till 1937. Her photograph lying on the couch (*Chaise Long*), which is their joint design, is left.

The 3rd person is Josephine Baker (1906-1975). She was a chanson singer and dancer from U.S., and got extreme popularity in Europe. She was a woman of the different type from Yvonne. The nude sketches of her slender body were left. They were drawn, in the ship (*Leticia*), in his return of the lecture tour in South America in 1929.

The above <female images> were related to some particular women. Their shapes were closely connected with particular and real women. So their deformation was not easy.

But other anonymous nude sketches existed, which did not converge to any specific woman. His return to the anonymous nude sketches after Purism could be traced back to 1925-6. He drew the naked dancers' sketches in the dressing room of a music hole. It was immediately after the farewell from Ozenfant at the beginning of 1925.

Some years later, he made anonymous nude sketches of prostitutes of Brazil, when he visited on the occasion of the South America lecture (1929) [5].

Finally the women who we should mention are those of Kasbah in Algiers, where he visit-

ed in 1931 for the urban design of Algiers. He made many anonymous nude sketches obstinately.

In his subsequent works, anonymous <female images> became indispensable elements. They were not connected with any particular woman. In other words, they were not accompanied by a specific outline. In other words they had no fixed outlines. Therefore they were allowed deformation. As a result the deformed images became to be immeasurable.

Probably the image was brewed inside him as autonomous image which was independent from the opposition to the male. It corresponds to what Deleuze calls *hecceité* (particularity and identical properties of itself) or *intensité* (intensity). Sawano has described *hecceité* as follows.

It is not pointing out “this one”, comparing “this one” with “that one”. When something of *intensité* is uniquely pointed out directly as “this one”, the individuality of the phenomenon is called *hecceité* [6].

Hecceité is the autonomous difference with neither oppositional objects nor comparison to other things. Furthermore, Sawano explains as follows.

Hecceité is what is immeasurable by the (suitable) scale or the degree and we miss the world image composed with *hecceité* just because we are confused by false measures or counterfeit degrees [6].

Immeasurability of *hecceité* is opposed to the geometry whose literal meaning is “metric (measuring or measuring method) of the earth” as mentioned above.

In summary, Le Corbusier's female image has another property. It is corresponded to *hecceité* by Deleuze. It is both independent from any comparison and immeasurable.

4. AUTONOMOUS METAMORPHOSIS FROM <OBJECT TYPE> TO <FEMALE IMAGES>

After 1928, <female images> became to be accompanied with <object type> or <object with poetic reaction>. The former had been the exclusive object of Purist painting, and the latter was introduced just before 1928.

How different genres of images coexist? To reply to this question we would like to propose the following hypothesis, i.e. their coexistence was made through the process of identification with deformation and formal association.

First he described on his formal association as follows. “When I saw the sketch of the picture of *still life* (1927-40) rotated 90 degrees, a series of *bull was born*” [7-8]. If it is true, we could assume another association. Such association with other objects, like the generation of image of a bull, was possible with deformation of <object type>. And this association would lead him to the transition to human body.

We would like to explain the process carefully. When a shape was drawn in a picture, it was related to a certain referent, for example, “a pitcher”. And it had a semantic attachment or meaning based on a fixed context. In Purist paintings, a “pitcher” meant the universality as <object type>. Because it was chosen in accordance with Purist rules whose model was laws of science.

However, one might look at the shape itself without the signification context. At that time, his interest was directed to the shape itself drawn in the painting. It is a moment to be able to call the beginning of an autonomous shape. In this level, the shape is played freely as a shape without its original meaning. It is a floating play like seeing various shapes in an amorphous cloud in the sky.

In the next step, an amorphous shape begins to change as a cloud. It is transferred to another shape of something else. Finally the transition fused two different images.

We would like to re-verify this hypothec process in a series of his paintings; from *Still*

Life (1926-7) to *Composition - Figure of person* or *Composition - logarithmic spiral* (1929) (Figure 1-5).

First *Still Life* (oil painting, 1926-7) is a transitional work from Jeanneret to Le Corbusier (Figure 1). Except the writing of “Φ” and signature of “L-C 25”, this painting did not significantly depart from Purist rules. But the shape of a pitcher, which was one of main objects of Purism, was slightly deformed. But it kept barely the typical profile (outline) of a pitcher as <object-type>.

Next is the immediately following sketch, *Still Life* (study, 1927), where the shape of a pitcher on the right was more deformed (Figure 2).



Figure 1: Jeanneret, *Still life*, 1926-7.

Jornod, N. et Jornod J.-P., *Le Corbusier catalogue raisonné de l'oeuvre peint Tome I*, Skira, 2005, p. 412, FLC 331.

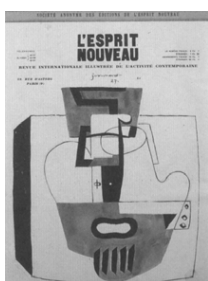


Figure 2: Jeanneret, *Still Life* (study), 1927.

Poma, M. C., ed., *Le Corbusier Pittore e Scultore*, Arnoldo Mondadori Editore, Milano, 1986, pl.47, FLC 1206.



Figure 3: Le Corbusier, *Still Life with books and egg stand*, 1928.

Jornod, N. et Jornod J.-P., *Le Corbusier catalogue raisonné de l'oeuvre peint Tome I*, Skira, 2005, p.443, FLC 339.



Figure 4: Le Corbusier, *Composition and head* (study), 1928.

Le Corbusier le dessin comme outil, Fage éditions, Lyon, 2006, cat. 69, FLC 216.

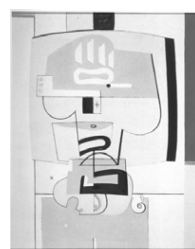


Figure 5: Le Corbusier. *Composition - Figure of person* or *Composition - Logarithmic spiral*, 1929.

Jornod, N. et Jornod J.-P., *Le Corbusier catalogue raisonné de l'oeuvre peint Tome I*, Skira 2005, p.465, FLC 171.

It shows asymmetric deformation as if it were inflated with internal pressure. It looks to be equilibration of the tensions of the inward force from outside and the outward force from inside. As a result, the original image of <object type>, which should keep the original outline as a universal object, disappeared.

Thirdly, *Still Life with books and egg stand* (oil paintings, 1928) closest to this study was made (Figure 3). It has the sign "Le Corbusier". The eggs and egg stand (in the upper part of the painting) and something like the legs of the table (in the lower part) were added. Texture (hatching), which had not been drawn in Jeanneret era, was specific not to Jeanneret but to Le Corbusier as a painter.

Fourthly, *Composition and head* (1928) is another immediately following study, where a profile of person's head was added on the top in (Figure 4). In this sketch the distorted outlines of the pitcher were overlapped with a human body.

Finally *Composition - Figure of person* or *Composition - Logarithmic spiral* (oil painting, 1929) was made from the above study (Figure 5).

This fusion of the different genres of objects was different from the idea of Purism (probably by Ozenfant). His idea was shown in the following description (*L'esprit Nouveau*, vol. 27, November, 1924).

We select our starting point as follows, that is normal conjugation without deformation which breaks the type (of objects) [9] (the author inserted parentheses).

Besides it was different from simple deviation from the rules of Purism as follows; to make a dice of a cube with adding eyes in the cube by Jeanneret. To make a dice of a cube was direct rejection or resistance to the rules of Purism (the prohibition of "dice"), but the fusion of different genres of objects was a more endogenous play of shapes.

In his past, La Chaux-de-Fonds era, he had discovered the geometric laws in the shapes of

natural objects and he had evolved them into decorative design. In the subsequent Purism painting, <object-type> kept universality within its constant outlines. Geometry was embedded there. On the contrary, natural objects in the paintings since 1928, i.e. <object with poetic reaction> did not have a geometric stylization like this. They did not have any fixed geometrical shape. When an <object type> was deformed, its outline might be fused with that of the <object with poetic reaction>. Then it would be released from the fixed real image. It was, as it were, the birth of <shapeless object>. In the next step, the ensemble of <shapeless objects> produced another association to a new object. Probably it was through this process that he made *Composition - Figure of person* or *Composition - Logarithmic spiral* (1929) (Figure 5).

5. <FEMALE IMAGES> AS GEOLOGY

He unified his activity as a painter using women late in life, which showed his continuous adoration of women. We would like to repeat his phrase.

I have drawn or painted only women, or the images, or the symbols, or the geology of women. From a tip of a foot to the top of the head there are stages and altitudes [1].

If he had drawn or painted only women, how should we consider Purist paintings? To reply to this mystery was the purpose of this paper. In order to

To reply it, we would like to make geology (ge-o-logy) oppose against geometry (ge-o-metry). The former literally means "logic of the earth", and the latter means "metric of the earth". If we accept this opposition, the female image corresponding to <ge-o-logy>= "logic of the earth" would be opposed to <ge-o-metry>= "metric of the earth".

This opposition also leads to the next corollary opposition of "straight lines and a right angle" with "curves". It was seen in the "pack-donkey's way and man's way" above

mentioned [4].

Furthermore <geology> would be immeasurable, if it is opposed to <geometry>= “metric of the earth”. Therefore <geology> of women connects to *hecceité* [6].

In summary we could explain his change as follows. The measurable <geometry> of Purism was transferred to unmeasurable <geology>. Deformation from Fig.1 to 5 suggests this transfer.

However, it is not a sufficient explanation. It lacks his motivation. So we would like to show other characteristics of Purist paintings related to his transition to <geology>.

Speaking the conclusion before, there exists double conversion of suppression. <Object type> selected in Purism is the first conversion of suppression. In Purism, sensation was divided into the first sense and the second sense. And a cube and a dice were compared. The first sense “seeing a dice as a cube” was universal because it was based on geometry. In contrast, the second sense “seeing a dice as a dice” was dependent on races and cultures. So it cannot be universal. And daily industrial commodities or musical instruments, called <object-type>, were selected by <mechanical selection> that follows the natural selection. It was something changed from suppressed <geometry>. In other words, the <object-type> was a conversion of suppression.

His deviation from Purism began with drawing a dice soon after the establishment of Purism. It was just before the exhibition of Independent Salon late in 1922. It led him to a chain of further updating rules. And geometry based on the first sense shifted from the center of gravity when he composed his painting.

However, as long as it is any update of the rules of Purism, he can not escape from the spell of the rules. On the contrary, the more he tries to deny the rules, the more he is restrained by the rules. To escape from the spell completely, in other words in order that Jeanneret became a completely independent painter as Le Corbusier, it was necessary to cut off the chain of updating Purist rules. Second conversion of

suppression was required. Therefore he had to revert to his origin which was opposed to Purist <geometry>. It was his return to women. And finally the women’s images unified his activity as an artist in conjunction with his transition from <object type> to the human body.

If we think like this, the re-introduction of the <female images> will not be able to be a simple denial of Purist painting. We should rather consider as follows. A new world of coexistence was created, where <female images> intervened between geology and geometry. Finally <geology> included <geometry>. The former corresponded to <female images> and the latter underlay the basis of suppression of Purism. The horizon let him say as follows. “I have drawn or painted only women...”

REFERENCES

- [1] Gallery Taisei ed., *Le Corbusier “Entre deux (Between two)” Exhibition of le Corbusier’s Prints*, gallery TAISEI, Tokyo, 1994.
- [2] *Le Corbusier Peintre avant le Purisme* (After Cubism). Musée des beaux-arts de La Chaux-de-Fonds, 1987.
- [3] Ozenfant, A. et Jeanneret C. E., *Le Purisme* (The Purism). *L’esprit nouveau*, 4, Éditions de L’esprit nouveau, 369-86, 1921.
- [4] Le Corbusier-Saunier, *Le chemin des ânes le chemin des homes* (Pack-donkey’s way and man’s way). *L’esprit nouveau*, 17, Éditions de L’esprit nouveau, 1922.
- [5] *Le Corbusier le dessin comme outil* (Le Corbusier the dessin as tool), Fage éditions, Lyon, 2006.
- [6] Sawano, M., *Leveraging Deleuze*, Keiryu-sya, Tokyo, 2009.
- [7] Jornod, N. et Jornod J.-P., *Le Corbusier catalogue raisonné de l’oeuvre peint Tome II*, Skira, Milan, 2005.
- [8] Moos (von), Stanislaus, *Le Corbusier Elements of a Synthesis*. The MIT Press, Cambridge, 1979.

- [9] Ozenfant, A. et Jeanneret, C. E., Idées personnelles (Personal ideas). *L'esprit nouveau*, 27, Éditions de L'esprit nouveau, sans pages. *La peinture moderne*, Les Éditions G. Crès et Cie., 1924.

ABOUT THE AUTHORS

Michio KATO, Dr. (Eng.), is a professor of the University of Tokyo. His research interests lie in architectural drawings, especially of Le Corbusier. He published the book titled, *Le Corbusier – Space and Time in His Architectural Drawings* (2011) and *The Birth of Le Corbusier as a General Artist- critic, painter, architect* (2012). His e-mail and postal address is as follows: kato@idea.c.u-tokyo.ac.jp. Department of General Systems Studies, Graduate School / College of Arts and Sciences, The University of Tokyo, 3-8-1 Komaba, Meguro-ku, 153-8902 Tokyo, Japan.

COLOR SCHEME SCRAPBOOK USING A CHARACTER COLOR PALETTE TEMPLATE

Ryuta MOTEGI¹, Yoshihisa KANEMATSU¹, Takahiro TSUCHIDA², Koji MIKAMI³ and
Kunio KONDO³

¹Tokyo Metropolitan University, Japan ²Riken, Japan ³Tokyo University of Technology, Japan

ABSTRACT: Character in visual content is important as well as the script when considering its importance and significance in making of value and charm or the production. We proposed a Design Engine[1] to improve efficiency of character development by providing support to the planning phase, where character blueprints are drawn based on literal information provided by the producer. Previous research exists on color scheme design work, but these methods are not useful for character design for producers and directors. This paper aims to provide a character color scheme support system that uses a scrapbook made up of the color palette of existing characters. This is one part of our Design Engine, whose purpose is to support the process of character making that is based only literal information. The concept of the Digital Scrapbook (DSB) is to set up a personal database that helps creators pin down the appearance of characters they want designers to sketch.

Keywords: CG, Character Design, Character Making, Color Scheme, Color Palette, Scrapbook.

1. INTRODUCTION

Video content, such as animations has become possible to easily produce by advanced CG technology. To produce the animation alone is not uncommon in recent years, and the video content production came to be widely performed. Characters are those that determine the value and appeal of the video content, along with the scenario. The color of the clothes and hair is one of the important elements that impress a character objectively in the video content. They are determined in many cases based on personality, setting of character and the background setting the world of video content [2]. Color scheme working on character is said that work sensuous, because it is dictated based on relevant as personality, setting and a variety of factors [3].

Color design is used in many fields related to daily life of space design and such as fashion design. There is a study, color image scale, color scheme system and such as harmony color scheme system. Kobayashi [4, 5] performed

application and impression evaluation of various designs by classifying the color scheme using the color image scale. Imahashi et al. [6] and Unami et al. [7] Were applied in the production of the illustration image that is convert the color harmony based on their color by extracting the image color representing characteristics of image. Cohen-Or [8] has proposed an image conversion using a harmonious color scheme. However, in those studies, there are three problems that have not been attention: for the use of the character design, conversion of the combination of the color scheme of character, the study of color conversion for each part of hair and clothes, cannot be found. In addition, the fashion design field also deals with the color scheme, but can't be used in combination of the color scheme of the entire character because clothes color and hair colors for animation character is not realistic. A purpose of this study is a search and management scrapbook creation at color scheme of existing character,

and to build a support system of color scheme for the new character. In this study, we propose a character standard template for collecting the color scheme of the existing character. And we propose a support method for character color creation and a search method using a color template. In this study, first, we create a character template of color scheme based on the results of analysis on every part of the hair, skin, clothes, on the color scheme of existing characters. Next stage, applied color scheme of existing character to color scheme template in order to make the data of the character color scheme. Then, we construct coloring support system that is able to extract the color scheme for creating own characters from a color combinations percentage of color scheme template. Finally, we conducted an evaluation experiment of the character color scheme using color scheme support system.

2. CONVENTIONAL APPROACH

In deciding the color scheme of character, it can make in the sense of personal, it is referring to the case in the past in many cases. Therefore, the current situation is highly dependent on the production experience and the sense of creator. Motegi et al. [1] Showed that it is possible to support the producer make a draft design of the character based on literal material using character design support system. However, it did not deal with color scheme support of character (Figure 1).

Wayama et al. [9] was classified a color scheme of animated characters. This study proposed a search system for making a comparison of color scheme by applying to template the face and clothes of the character. Thus, it has become possible to search for the color scheme of character with a personality and a main color used. However, color scheme generator and change process of character based on search results is not addressed (Figure 2). Therefore, performance enhancing of data library and,

function's implementation of color scheme simulate has been desired for color scheme support of character. This study constructs a digital-scrapbook of an existing character, and color search simulation system in order to solve this problem.

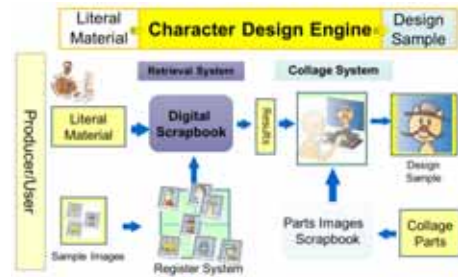


Figure 1: Overview of character design engine.



Figure 2: Classification of Character Color Scheme.

3. COLOR SCHEME TEMPLATE AND ANALYSIS OF CHARACTER PARTS

3.1 Parts quantity of character

We were classified hair, skin and clothes, of the character, in every part for simulation of color scheme. Character's hair has a hair of various colors unlike a real person. Also, we think that represents a strong personality of the character. The color of skin is no large difference in each character, because it's made of the human deformed. And clothes, color scheme and how to

wear was variable. Therefore, we examined the 51 characters in order to know the number of types of ornaments and clothes, of the character needed for color simulation. In this study, clothes are able to wear as pants and shirts, ornaments are able to put on as tie and ribbon. And, carry out analysis with the addition of hair and skin.

Investigation result showed that the ornaments are formed by the 0-2 type, the clothes are formed by the 1-3 type, as shown in Table 1. Combination of ornaments and clothes, of the all characters of other than two characters has subsided in four types within. Further, hair and skin are composed of one color basically, ornaments and clothing are constituted a plurality of colors, a plurality of parts, basically. From the above results, in this study, we deal in the color scheme of the character based on one type, hair color, one type skin color, three type clothes color and two type ornament color.

Table 1: Parts quantity of clothes and ornaments, of character.

		Ornaments				total
		0	1	2	3	
Clothes	1	4	5	8		17
	2	11	4	4	1	20
	3	4	9	1		14
	total	19	18	13	1	51

3.2 Character color scheme template

Here, we describe the character templates that can be informed intuitively the image of the character color. For this purpose, we examined the parts that are used in ornaments and clothes of an existing character. We analyzed the area of clothing parts to wide variety that are originally designed from the clothes of existing school uniform, suit, white robe, like Kimono. As a result, Hair, Neck, Shoulder, One-third sleeves, Half sleeve, Three-quarter

sleeves, Long sleeves, Chest, Waist, Hip, Inseam, Above knee, Knee, Ankle, Shoes, Shoes destination, Coat, Tie, so on divided 58 areas by 18 solid lines. The color template was created as intuitively understandable placement of the color of the divided parts, each region, such as clothing, as shown in Figure 3. Figure 4 shows an example in which you have registered with a scrapbook of character color scheme.



Figure 3: Character color scheme template.



Figure 4: Example of registering character color scheme scrapbook

4. PROPOSED SYSTEM

4.1 Overview of color scheme support system

Figure 5 shows an outline of color scheme support system. Color scheme support system is composed in two search system and reflection systems and registration system. Figure 6-7 shows the steps of the color scheme simulation. First, the color scheme of existing character registers as a color scheme image to color balance for each part hair color, skin color, clothes color. And, to set the color scheme area of each part hair, skin, and clothing against line drawing, and reflect the color balance and color scheme to select a color scheme image that has been searched. After that, fine-tune of the color scheme to each part, and export as image data.

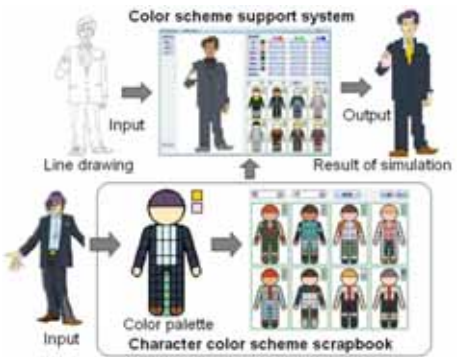


Figure 5: Color Scheme Support System with DSB

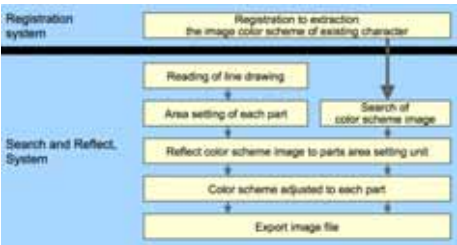


Figure 6: A steps of registration, search, simulation, of character color scheme.

4.2 Registration and character color scheme scrapbook

Figure 8 shows a flow of the character color scheme registration. Figure 9 shows a color scheme registration system. The registration system constructs a library of existing character data. A loading the image of an existing character, then extract 7 type base color of total as 3-type of hair color-skin color-clothe color, and ornaments 2 type. After extraction, to set 58 locations color of the total to the template that you created while compared to the parts placement of an existing character. The save of data when the exported as CSV data, and it is stored in the character color scrapbook. The contents of the data to be saved is seven parts of color and, color information about the 58 area template setting.



Figure 7: Step of color scheme processing.

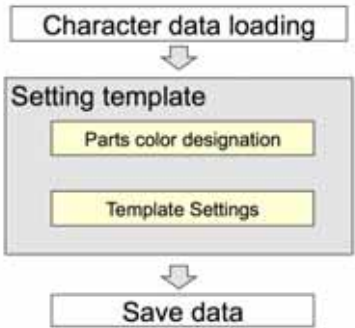


Figure 8: Flow of existing character color scheme registration.



Figure 9: Color scheme registration system of existing character.

4.3 Search system

Figure 10 shows a flow of character color scheme search. Then, figure 11 shows a color scheme search system. Reflection and Search system is performing system of color scheme to design draft of created to producer on the basis of the character data registered by the registration system.

First, load the image of the design draft of uncolored, and to setting the color area on seven type base color of total as hair, skin, clothes, three types, and ornaments of two types. Searches from the image of hue and parts of character registration data, and reflect in color area of seven types.

The Make minor adjustments by using the color picker and color bar from the data that reflects, save as image data the finished image, and it is color scheme draft. Further, if saving data is possible to perform the re-edit, to export data containing color data as a CSV data. Figure 12-14 shows the state of each step.

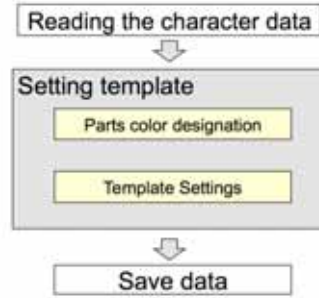


Figure 10: Color scheme registration system of existing character.



Figure 11: Color scheme search system (character line draw input)



Figure 12: Part area designation.

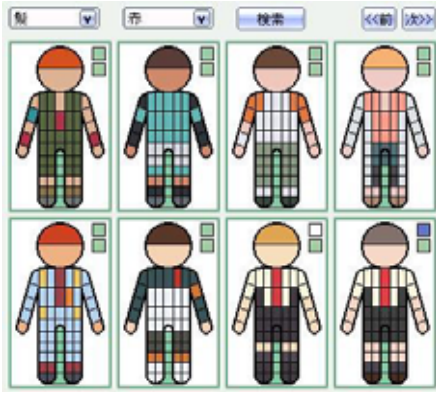


Figure 13: Search Results of character color scheme.



Figure 14: Application of color scheme results to character image.

4.4 Color scheme experiment

Figure 15 and figure 16, both show the result that was fitted to the line draw of the seating area of each part from hair, skin, clothes, etc., the color image of the existing character. If it were fitted with a color corresponding to the clothes, it is descending order of the color area. Therefore, Figure 16 shows a case where the result of the color area of the shirt and coat is different from the template.

Comparison of the character color scheme is easily on the basis of line draw, it was possible to produce an intended color.

5. EVALUATION EXPERIMENT AND CONSIDERATION

15 persons took test subject for system evaluation. Figure 17 shows an example of color scheme creation results using a color scheme support system by test subjects. Table 2 shows the results of surveys made to test subjects. Average value (5-point scale) of right the table shows that is a useful higher the number to the feature. Figure 18 shows a “7. Color scheme simulation (graph left)”, “8. Intentional color management (graph right)”, the 2 items of the color scheme simulation of the entire system was 4.5. These account for more than 85% of the total in each evaluation "very good" and "good" respectively. In addition, improvement of the color scheme search algorithm and increasing the number of registered color scrapbook is desired. The results of the evaluation experiment are:

1. By setting the part area of the character, it is possible to easily create a character color scrapbook capable of color conversion.
2. By reflecting the character line drawing color scheme of an existing character, it is possible to intuitively consider a variety of character color scheme.
3. Production support of Intentional color is possible by using the proposed system, and reduces the number of steps of the character color design work.

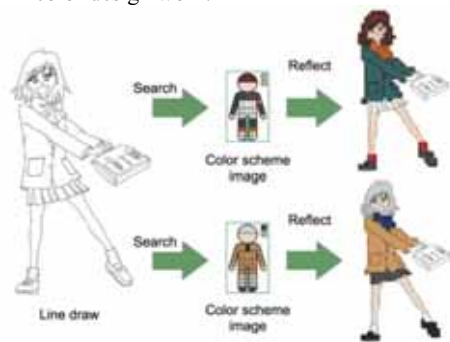


Figure 15: Part 1 of the color scheme experimental results.

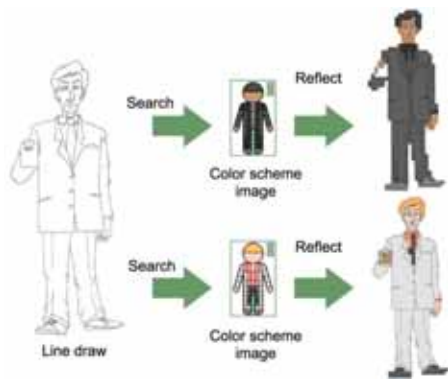


Figure 16: Part 2 of the color scheme experimental results.



Figure 17: Example results of color scheme experiment by the test subject.

Table 2: Result of Evaluation experiment.

Question contents	5	4	3	2	1	(M)
1. Setting of parts area	3	11	1	0	0	4.1
2. The search result display	2	7	5	1	0	3.8
3. The fit to the design image	7	5	2	1	0	4.2
4. Color scheme adjustment	8	6	1	0	0	4.5
5. Flow of system use	7	7	1	0	0	4.4
6. Steps of entire system	7	8	0	0	0	4.5
7. Color scheme simulation	10	4	0	1	0	4.5
8. Intentional color management	9	4	2	0	0	4.5

5= Very Good 4= Good 3= Neither 2= Bad 1= Very Bad

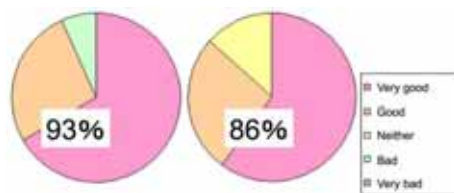


Figure 18: Example results of color scheme experiment by the test subject.

6. CONCLUSION

In this study, to create the character color scheme scrapbook that allows you to manage and search the color scheme of existing characters. And, we aimed to build a support system color scheme for created the character. For this purpose, we had proposed a character standard template for collecting the color scheme of the existing character. Then, we constructed a color scheme support system for performing the color of character entire, and the search method at Character color scheme scrapbook using the template.

As an experiment result:

1. Rather than performing a color from the beginning of the uncolored design image, the color processing system proposed in this study is more efficient.
2. Character color scrapbook can increase the efficiency of using the existing character color scheme

The future works are:

1. By registering the color image with more character color scrapbook is necessary. Thereby, Color work will be able to more smoothly make to broaden the simulation.
2. To achieve the high performance of the search method of the Intentional color scheme by the system user.
3. In order to correspond to the more detailed design, and can be set more finely in any setting area of each part of the grid.

REFERENCES

- [1] Ryuta Motegi, Ryoichi Matsumoto, Kunio Kondo, Mitsuru Kaneko, A Proposal of Designing Method Based on Literal Materials, *proc. of NICOGRAPH*, 2007
- [2] Mitsuru Kaneko. Making of Video Content, The basics of Content Engineering. *Born Digital, inc.* 2007.
- [3] Yasuko Shibaguchi and Michiyo Yasuda. Animēshon No Iro Shokunin. *Tokuma Shoten Publishing Co., Ltd.*, 1997.
- [4] Shigenobu Kobayashi. Color image work. *Kodansha Ltd.*, 1995.
- [5] Shigenobu Kobayashi. Color image scale revision Version2. *Kodansha Ltd.*, 2001.
- [6] Hiroki Imahashi, Kunio Kondo, Yoshiaki Machida, Masahiro Takahashi, Knowledge Based Color Coordinate System and its Application, *International Journal of ADADA, Vol.1*, pp. 37-42, 2004 (proc. of ADADA 2003)
- [7] Yukiko Unami, Kunio Kondo, Area Segmentation and Color Exaggeration for Image Synthesis, *International Journal of Asia Digital Art and Design Association, Vo.5*, pp.38-43, 2006.12
- [8] Daniel Cohen-Or, Olga Sorkine, Ran Gal, Tommer Leyvand, Ying-Qing Xu, Color Harmonization, *ACM Transactions on Graphics (Proceedings of ACM SIGGRAPH 2006), Vol.25, No.3*, pp.624-630, 2006
- [9] Shunji Wayama, Mitsuru Kaneko, Research to classify the color scheme of Japanese anime character, *NICOGRAPH Spring Festival in TAF*, 2007

THE AUTHORS

- 1. Ryuta Moteg, Faculty of System Design, Tokyo Metropolitan University, Japan.
- 2. Yoshihisa Kanematsu, Tokyo Metropolitan University, Japan.
- 3. Takahiro Tsuchida, Riken.
- 4. Koji Mikami, School of Media Science, Tokyo University of Technology, Japan.
- 5. Kunio Kondo, School of Media Science, Tokyo University of Technology, Japan.

COMPLEXITY OF PUBLIC SPACES SYSTEM BETWEEN KEY TALL BUILDINGS IN CITY OF SZCZECIN GEOMETRICAL ASPECT OF PUBLIC SPACES IN 3D CITY MODEL

Adam ZWOLIŃSKI

West Pomeranian University of Technology, CUC, Szczecin, Poland

ABSTRACT: The paper is strongly established on compilation of urban analyses of public spaces with environment of 3D city models. It is based on the understanding, that urban space is a composition of “positive” and “negative” – physical built-up environment and non-permanent dynamic pattern of users and their activities. Presently, it is more and more proved, that public spaces are the driving force and determinants of urban quality in most cities. The most important public spaces are commonly located within city center areas, however very often overlapping with locations of tall buildings. These are either tall historical buildings of cathedrals, churches, towers, palaces etc., or modern skyscrapers. The tall buildings are not the case itself for this article, but they are used only as a pretext for choice of case study sample areas located nearby city “markers”. The 400.000 inhabitants city of Szczecin in north-west part of Poland has both of the mentioned “markers” in the city center. There is a lot more European cities where key tall buildings are located nearby or within important public spaces (squares, streets, boulevards etc.). Recent period of rapid development in computer technologies affects also sphere of architecture and urban design. Technology of 3D city models is being intensively developed for different uses and also it seems to be very perspective and attractive for application in terms of advanced urban analyses. It inspires a multi-purpose approach to analysis of cityscape using 3D models, particularly a “negative” part – called 3D void between buildings. The paper uses concept of representation of public spaces in 3D city model as 3D multi-face bodies with measurable parameters. This gives an opportunity of converting typical 2D urban analyses into the world of advanced spatial geometry including more parameters describing 3D solids. The attempt of parameterization of public spaces using concept of 3D solids representing space between buildings will be a subject of this paper. The 3D city model will be used to generate 3D geometry of public spaces in the selected area. The detection will be processed using generated point cloud and cut surfaces in general 3D model. The general outcome will be multi-aspect observation of process of generating 3D voids in specific selected areas of city using 3D city models.

Keywords: 3D virtual city models, public spaces, urban analysis, tall buildings

1. CONTEXT

The main determinant of interest and specific approach to problems of space between buildings is observation of contemporary cities, where urban planning and design meets rapid technological development in sphere of computer tools and techniques. Intensive growth and transformation of cities can be observed

also in urban projects focused on public spaces as important urban landmarks of cities. On the other hand, there is a growing potential of computer-based tools in present age of information, high technologies, optimization trends and growing importance of open-source data available for different applications. It is also about the growing applicability of 3D city models, not only for commercial, presentation

use, but also for analytic applications. Geometrical and numeric interpretation of common phenomena, also urban areas – cities, is happening right now. The real invisible database of city life is written in system of public spaces. The crucial interest of author is how the life between buildings appears in different types of spatial forms, volumes and specific geometry of form. The approach is resultant between general analytic approach based on classic observation and understanding of life between buildings by Jan Gehl [3], and technologically advanced, analytic and computer-based approach of Space Syntax. It also refers to notion of spatial configuration and spatial cognition introduced by Kevin Lynch (1960) [10], also undertaken by Bill Hillier (1984) [4,5] and A. Penn [11] - Space Syntax. The relation between spatial configuration and spatial cognition is also widely discussed by Y. O. Kim (2001) [7]. The author's work realized within research team of CUC [15] is focused on application of 3D city models in advanced urban analyses. Author's intention of introducing analytic method in environment of 3D city models is decomposition of subconscious scheme of visual interpretation of space between buildings, towards objective listing of parameters – then reading again city database using intended algorithms. New computer tools applied in 3D city models environment allow to introduce multi-aspect urban analyses (typology, geometric complexity, capacity, quantitative data etc.). Introducing notion of complexity refers to assessing geometric parameters of space between built-up structure of city. At the present stage scientific research is focused on physical characteristics of public spaces and optimization of 3D city models as environment for advanced urban analyses.

1.1 Public spaces: relation to tall buildings

Public spaces, as a part of general cityscape are the main interest of author, but the ongoing CUC 2Tall project, focused on general impact of tall buildings on cityscape, is also oriented towards observation of public spaces

surrounding of in direct relation to tall buildings. The observation is very interesting because localization and geometrical characteristics of volume between / around tall buildings is very specific and directly dependent on scale and quantity of such buildings. One of the observed typologies of tall buildings in European cities is simple division on historical tall buildings located in core historical areas of cities (these are mainly characteristic heritage sacral buildings with towers dominating over old urban structures), and modern developments of high-rise buildings in different type of localizations.



Figure 1: Public space beneath tall building. La Defense – Paris. Photo by K. Czyńska.

New tall buildings are often prestigious and attractive architectural forms. On one side, the impact on cityscape is important factor, but on the other, public spaces beneath create the link to the city. Tall buildings are often located in key areas of city centre and their direct surrounding is very important for system of public spaces. According to Czyńska [2], there are different types of concentration and localization of tall buildings. We can observe spatial concentration (as it is in Vienna), dispersion of structure (Köln) or visual concentration (Frankfurt). The case of Szczecin is more towards dispersion of structure – but limited to basically two tall buildings in scale of the city. However the buildings are located in key areas of city centre – the Radisson / PŻM building complex site analyzed in application part of paper is located at one of important public

spaces axis of Szczecin. Detailed description of the project was presented in 2013 by P. Rubiniowicz [16].

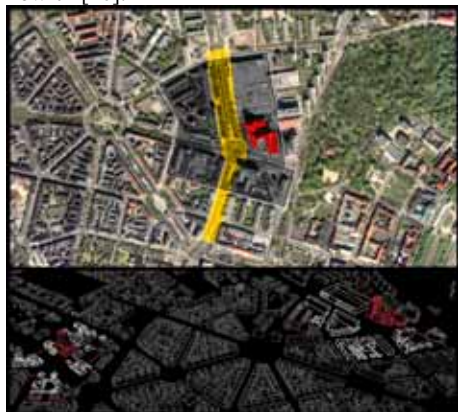


Figure 2: Localization of Radisson / PŻM building complex – Szczecin

1.2 Public spaces: significance and approach

This paper refers only to the physical and geometrical aspect of wider and very complex phenomena of public spaces in cities. The importance for sustainable development of cities is undisputed. Quality of public spaces is one determinants of general quality of cityscapes. The space between buildings is still unexplored database to be discovered using environment of 3D city models and, following rapid technological development, to describe characteristics of the space in parametric, comparable and analyzable way. Significance of space between buildings lays also in potential of anticipation of future activities and spatial organization. Moreover, advanced 3D technologies facilitate better understanding of public space as abstract geometrical volume limited by physical elements in urban space

2. PUBLIC SPACE AND 3D CITY MODEL

General understanding of public space can be easily divided into two important approaches. One is urban-oriented approach treating public space as general open space between built-up urban blocks, second is user-oriented

approach understanding public spaces as specific, legible areas in city where intensive public life occurs. Popularity of 3D models of cities is rapidly growing in last decade. Google Earth opens possibilities to cover with 3D models practically all cities worldwide. There are also other standards, such as CityGML [8], creating powerful potential of adaptive use of 3D city models for advanced urban analyses.



Figure 3: Szczecin 3D city model with. Tall buildings in red.

The other important issue about public spaces and 3D city models is, that the models are mainly developed for built-up areas, the left part is still insufficiently explored.

2.1 3D city models and techniques

Combination of rapidly growing computer capacity and development of 3D city models can result in a powerful open-source tools for multifunctional applications. However, there are still many difficulties behind. One of the observed problems is standardization and optimization to diminish differences in available formats of data. The other technical problem is visualization and processing of big data included in 3D city models.

```

2393,72; -681,84; 43,69 Select objects: LWPOLYLINE Layer: "site2"
2443,7; -681,84; 27,49 Space: Model space
2384,37; -682,76; 27,49 Color: 1 (red) Linetype: "BYLAYER"
2391,45; -682,63; 43,69 Thickness = 63.5000
2210,02; -682,51; 28,35 Handle = FOA
2231,47; -690,6; 28,35 Closed
2219,58; -687,06; 28,35 Thickness 63.5000
2224,84; -688,64; 28,35 Constant width 0.0000
2219,23; -683,74; 28,35 area 88.4733
2261,66; -686,59; 28,58 perimeter 52.6230
2259,92; -685,48; 28,58
2252,58; -684,24; 28,58 at point X=1381.0954 Y=-180.3242 Z=29.8900
2258,39; -698,42; 28,58 at point X=1379.5182 Y=-186.3243 Z=29.8900
2259,09; -695,87; 28,58 at point X=1387.4905 Y=-199.8889 Z=29.8900
2249,49; -695,84; 28,58 at point X=1393.5000 Y=-201.4300 Z=29.8900

```

Figure 4: Example extract from numeric image of 3D city.

Moreover, important factor related to data included in 3D models is so called LOD (Level of Detail). LOD is extremely important to get comparable data from city models developed on different platforms – different accuracy is demanded for different analytic applications. Representation of public spaces is specific problem in 3D models, because of complexity of small scale detail – very resource-consuming for processing in computers.

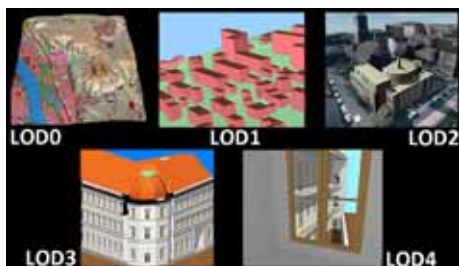


Figure 5: Five levels of detail defined in CityGML standard. Figure by T. H. Kolbe, G. Groeger, L. Pluemer [9]

All the existing formats of city models (Google, CityGML, game engines, ESRI etc.) contain database of parameters useful for analytic interpretation. There is also potential of GIS predefined tools (such as Quantum GIS with Grass etc.) to be applied on standardized data content. Finally there is also a bit other branch of tools for simulation purposes, particularly interesting in sense of Agent-Based Modelling for analysis public space use. The presented work developed within CUC is realized in combination of platforms: AutoCAD with individually programmed AutoLISP scripts (P. Rubinowicz) and Quantum GIS with Grass extension. The key and most important issue for all the presented data analysis is concept of writing 3D city and public spaces as a list of numbers and configuration of XYZ points in space (Figure 4). The odd, at the first glance, understanding the cityscape as a basic code or list of numbers, is a very important to get objective parametric data for processing with individually set up algorithms.

2.2 3D city model: public space samples

Selection of sample public spaces within city of Szczecin was driven by spatial characteristics of sites and assumptions of the ongoing 2TaLL project: “Application of 3D Virtual City Models in Urban Analyses of Tall Buildings” – realized by CUC research team at WPUT Szczecin. Both selected sites are located in the very city centre area of the city. Both are located directly at the only two modern tall buildings scattered in Szczecin. The entire 3D virtual city model of Szczecin used for sample application in this article was developed by CUC. The samples of Berlin 3D city model are used by courtesy of VCS Berlin – specialist firm for CityGML standard.

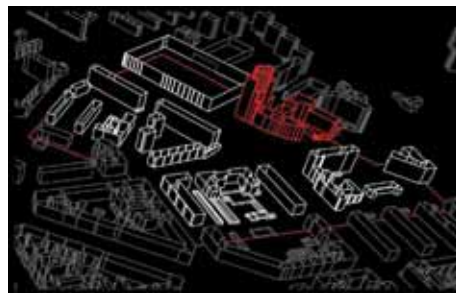


Figure 6: Selected site 01: Area at Radisson / PZM building complex.

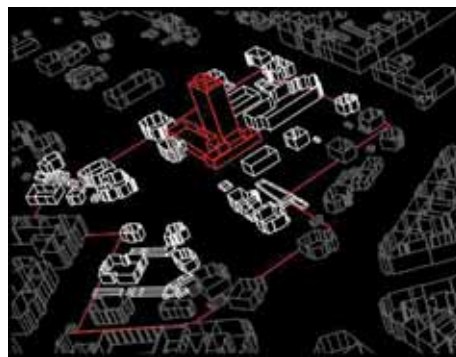


Figure 7: Selected site 02: area at local TV tower.

According to 3D geometry of selected areas, both represent type or varied urban built-up

area around single tall building adjacent to main city street grid. Both sites contain variety of building heights and sizes. Site 01 – mostly used for analytic sample application in this paper – has more irregular spatial shape and higher complexity in sense of variety of architectural forms surrounding the area.

3. 3D NEGATIVE OF CITY SPACE

Using environment of 3D city models for parameterization of space between buildings, in sense of its geometrical parameters and complexity, bases on general concept that urban space is a composition of “positive” and “negative”. The positive part is all the built-up area (buildings, structures etc.), the negative is the entire in-between space – introduced in this paper as the 3D VOID. Both of them, together with dynamic patterns and flow of users, contribute to the cityscape as a whole. It is extremely important to observe, that apart from the subconscious perception of physical elements in city as geometrical forms of particular dimension and volume, the space between physical objects has in fact exactly the same physical features – unseen, but possible to describe in parametric way. The exception in the presented approach is greenery – which is, in this particular case, not considered as a part of positive or negative in 3D city model.

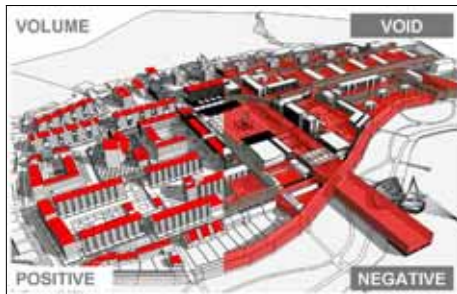


Figure 8: 3D city model: positive/negative concept of cityscape.

Understanding the “negative” part of cityscape as a hidden geometry is a key to introduce 3D geometrical interpretation of the volume li-

mited by physical elements of a cityscape. The environment of 3D city models has a potential to discover and explore this area in parametric way.

3.1 Introducing 3D VOID

The observed need of professional and advanced computer-based tools supporting analysis, planning and simulation of key urban processes occurring in cities as well as following tendencies numeric interpretation of common phenomena resulted in using geometry of 3D city model to create and process the invisible space between buildings. The volume of open space limited by built-up structures has been interpreted as a geometrical spatial form and individually called 3D VOID. This three-dimensional void is possible for geometrical tracing based on limitations of surrounding objects in 3D city model. This objects are: buildings surrounding open space (side limitation) – mainly modeled as 3D solids, terrain (bottom limitation) – modeled as 3D mesh and tops of built structures surrounding public space (top limitation) – represented in 3D model as top points of facades, roofs etc. The 3D VOID is in this context only static interpretation of physical environment of the city but in advanced way could be used for combination with data from user-oriented analyses of pedestrian behaviors and flows. The sample application of 3D VOID presents process of individual process and tool for automation of creating such object in 3D.



Figure 9: Conceptual visualization of 3D VOID idea in 3D city model.

The process of 3D VOID generation bases on simple geometrical rules and elements applied in AutoCAD / GIS platforms. The sample of automation was tested on three different city models (Szczecin, Berlin, Lublin) and proved reasonability of getting objective data for analysis with automated and optimized city models.

3.2 Generating 3D VOID

The very first step in process of generation of 3D VOID in 3d city model is marking analysis area and decomposition of the model into primitive geometrical shapes - lines and points. The 3D objects representing built-up structures (buildings) in city models are constructed in different ways (extruded polylines, meshes, solids etc.). For the purpose of optimization, AutoLISP procedure exploding model into horizontal / vertical lines and points was developed and tested.

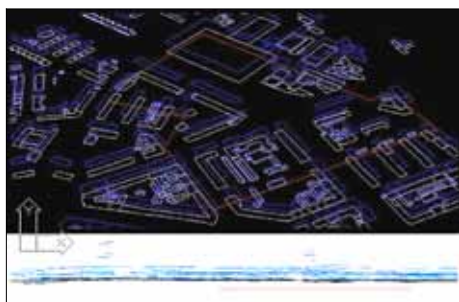


Figure 10: Fragmentation of 3D geometry into simple shapes.

Next stage refers to coordinates hidden behind all the primitive geometric shapes. The key for generation of 3D VOID is optimization of primitive geometrical content and automated generation of point cloud. The automatically generated XYZ points in 3D model for all nodes of exploded lines are the purest numeric code representing elements in 3D model. Figure 10 shows part of 3D model exploded in simple geometrical forms of lines / polylines. The numeric data is simple to export in text format, which is readable for major part of computer platforms and software (CAD, GIS). In the fol-

lowing stage the city will turn into points – only XYZ points.



Figure 11: Fragmentation of 3D geometry into simple shapes.

The structure of generated point cloud is that there are nodes representing bottom plan of buildings and structures, but also those representing top points. They differ in coordinates on Z axis, but in plan XY view the nodes have the same coordinates (Figure 11). Together with automated generation of a point cloud (they can be represented as any graphic symbol), the programming procedure makes identification of point cloud layers analyzing and dividing nodes into bottom – ground layer and top – roof layer.

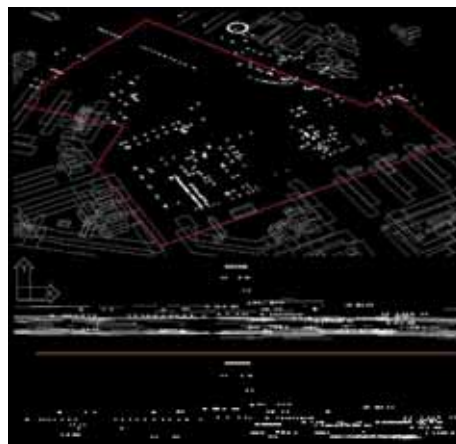


Figure 12: Automated generation of 3D point cloud with spectral view.

Figure 12 shows generated point cloud within

the limited area (red line) in perspective and spectral view. The distribution of points on Z axis is clearly visible in spectral view. Using simple operations like hiding, deleting or specific selection it is possible to export numeric .TXT data on decomposed 3D city model. It is also possible to use other algorithms to select specific points and identify other elements of city. The resultant data from point cloud generation is a clear and objective interpretation of the city in 3D environment. The notions of buildings, streets and other common elements disappear and are replaced by pure XYZ code, or extended code with more parameters (Figure 4). In the next stage of data processing there is a transfer of numeric data to other environment. The selected sample site of Szczecin was exported as a numeric point cloud to GIS platform – to Quantum GIS software. It is a powerful GIS computer tool having possibilities to use predefined tools of such data processing. The data processing can take place in core platform of QGIS, but also in its free analytic extension – Grass. It offers number of predefined analyses – as well as geometrical. One of them is very important to implement next stage of generation of 3D VOID.

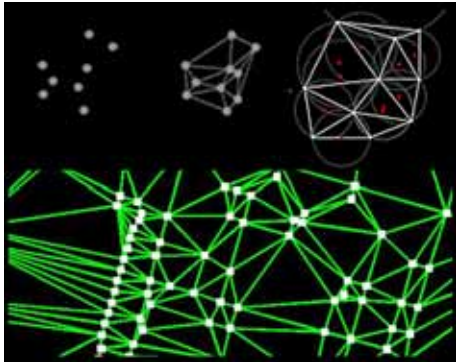


Figure 13: Delaunay triangulation of point cloud in GIS environment.

The point cloud representing top points of elements in 3D models (points of top elements of buildings) is converted into triangular surface

using Delaunay triangulation [15]. Figures [13,14] show general principles of Delaunay triangulation and process of triangulation executed on point cloud generated for area around key Radisson / PŻM tall building in Szczecin.

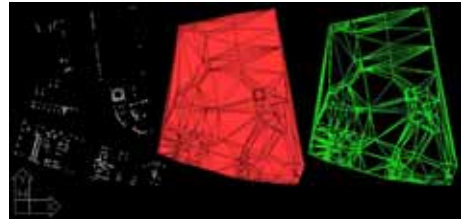


Figure 14: Flat surface resultant by Delaunay triangulation of point cloud.

The important feature of generated on GIS platform triangular surface is possibility of easy export to common and standardized .DXF format. It facilitates re-import of geometry to AutoCAD environment. The other specific feature of generated triangular surface is that the triangulation occurs in two dimensions, so the surface ready for re-import to CAD is basically flat (Figure 14). The following phase of generation of 3D VOID uses combination of GIS output with core 3D model CAD data.



Figure 15: Combination of GIS / CAD data in 3D city model.

Figure 15 presents combination of re-imported data from QGIS platform and core CAD content of 3D city model to prepared for last steps of generation of 3D VOID between buildings. The original point cloud is in XY coordinates exactly in the same position as imported trian-

gular flat surface from GIS. The next step is to get 3D triangular surface representing top surface of elements in city model (nodes with highest Z coordinate). Such prepared combination of data can be used for automated (by another AutoLISP script) extrusion of flat triangular surface to height (Z coordinates) of original point cloud. Figure 16 shows automated generation of 3D triangular mesh adjusting flat Delaunay triangulation to 3D point cloud executed in AutoCAD. This step is extremely important in sense of automation of the process of generation 3D VOIDS. It bases on individually programmed script for such 3D adjustments.

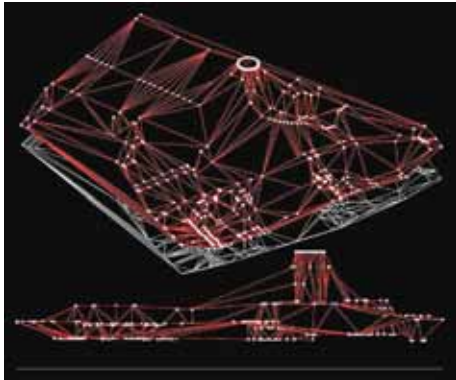


Figure 16: Adjusting 2D triangulation to 3D point cloud.

After having generated thin triangular 3D surface, it has to be extruded down to the ground level of model to get 3D solid in the 3D city model. The apparent 3D form is a new, alternative representation of city in 3D model. Top of the surface bases on connection between points on the edges of built-up structures. Additionally it is worth mentioning that also for exploring issue of complexity in 3D models, a special procedure was programmed for automated generation of links between points peer-to-peer each other, and will be tested soon. It gives possibility to analyze complexity of space between buildings in sense of number of possible connections. Getting back to the core city mod-

el, the following step assumes generating 2D plan from 3D city model (also automated by AutoLISP procedure). The 2D plan is used to extrude solids representing built-up volume. The extrusion can be executed to any height \geq than top points of structures in the 3D city model.

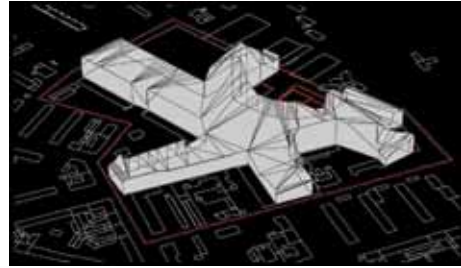


Figure 17: Pure 3D VOID after Boolean operations in 3D city model.

Finally the 3D VOID generation process needs application of another geometrical operations predefined in CAD software – popular Boolean operations. 3D Boolean operations are executed to subtract existing 3D structures from general triangular solid and finally to generate the entire 3D VOID. The last stage of ends up with automated generation of 3D general footprint of system of public spaces (space between buildings). The 3D VOID contains all the open spaces within selected area (Figure 18)

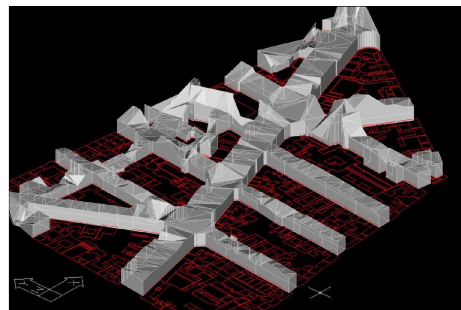


Figure 18: Sample 3D void of larger area in Berlin. Model by courtesy of VCS Berlin.

Final decomposition of 3D city model into particular, specific voids opens a possibility of

multi-aspect analysis of the space between buildings. It is also available to use generated 3D VOIDS as separate elements of independent urban analyses. The last important factor to understand is that each 3D VOID as geometric solid is also constructed and described by faces/vertexes and solids (Figure 19). All of them can be analyzed separately using different specific algorithms.

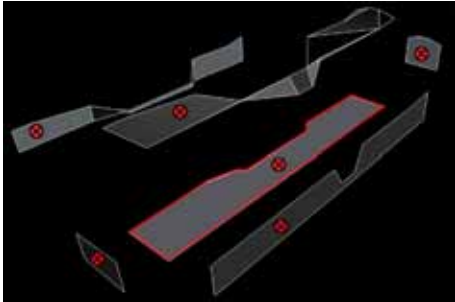


Figure 19: Theoretical construction of generated 3D VOID.

4. INTERPRETATION

All the presented stages of generating so called 3D VOIDS in city model result in intermediate data which can be read in parametric way. All of them: 3D linear structures, 3D point clouds, 3D triangular surfaces or solid 3D VOID itself. This part of the paper includes a brief introduction to selected possible 3 levels of interpretation of data in sense of geometric complexity of public spaces measured by 3D VOID method and tool used in environment of 3D city models. The interpretation will be presented for general point cloud, cut surface image of public space and 3D solid. The strength of the presented method is concept of incorporating potential of multi-environment universal data to be applied for advanced urban analyses. The components and intermediate products of 3D VOID method are applicable for analyses in both CAD and GIS environments. Particularly, presented QGIS platform has a great number of predefined tools to be used on the data. The interpretation of complexity of public space in

the sample sites was also about finding geometric regularities and typology of public spaces in cities. The 3D VOID method can deliver objective and comparable data for studies on specific public spaces in geometrical sense. There is a lot of parameters in the space between buildings represented in form of 3D object. They can be explored in different type of interpretations. The following part introduces some of them.

4.1 Point cloud

The 3D point cloud generated at the early stage of presented analytic process is a powerful set of data on general cityscape, as well as space between buildings. Figure 20 presents perspective view which contains potential of visual intuitive interpretation spaces in-between built-up areas. Firstly the concentration of points in space clearly indicates open structures of the city, then using proper algorithm to analyze and group particular distances between points we can refer to density of buildings. Also programming points with data on functions of particular parts of buildings it is possible to generate functional and user accessibility point cloud for public spaces. The point cloud has potential of comparable analysis of cityscapes in aspect of system of public spaces in entire city extent. Geometrical regularities of distribution of points is also readable parameter for typologies of public spaces around the city.



Figure 20: Sample 3D point cloud of area in Berlin. Model by courtesy of VCS Berlin.

The second extremely interesting mean of interpretation of point cloud is its effect of spec-

trum - using point cloud as a spectral view on system of public spaces. The elevation view of point cloud creates a spectrum of concentration of points showing geometrical regularity of city profile in both X/Y and Z directions. It also refers to concept of X-RAY observation of the city introduced in earlier research by CUC (K. Czyńska / W. Marzęcki / P. Rubinowicz). The presented on Figure 21 spectrum of point cloud also gives potential of observation and interpretation of flow of points at ground layer (Z=0 points) and top layer (building height Z points).

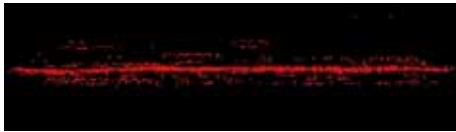


Figure 21: Spectrum of generated 3D city point cloud in elevation view.

The point cloud interpretation level is mostly intended for city-wide scale observations and analyses of space between buildings.

4.2 Cut surface image

The more detailed level of interpretation for both point cloud and entire 3D VOID is using sequences of repeatable offset sections to generate different patterns of public spaces in sense of geometrical characteristics of plan view from each cut surface. Figure 22 presents 3D concept of cut surfaces in 3D model.

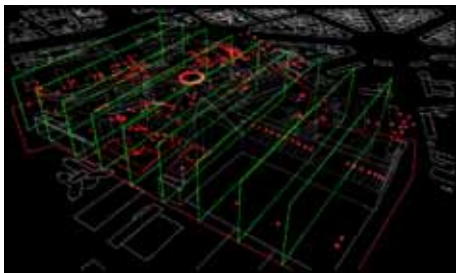


Figure 22: 3D cut surface interpretation concept.

Once, this level of interpretation focuses on identification of public spaces using point

cloud for reading typology created by cut surface analysis. The interpretation can be recorded as a common shape of polyline linking properly selected points in point cloud. The set of linear images has further potential of image data content analysis comparing the shapes and searching for regularities and typologies. The cut surface images can be also analyzed in sense of general proportion of open space to built-up areas in cities. (Figure 23).

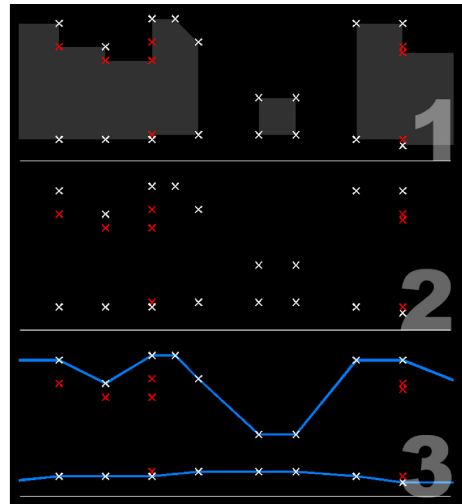


Figure 23: 3D cut surface complexity analysis of public space.

The cut surface level of analysis gives also more detailed scope on geometric shape and proportions of each public spaces itself as it is presented on Figure 24. by combination of linear scheme and solid objects in 3D model.

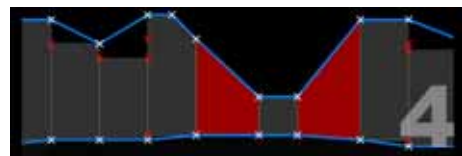


Figure 24: 3D cut surface complexity analysis – geometrical typology of public spaces.

The cut surface image level of interpretation of 3D VOID is extremely important for exploring phenomena of typologies in both scales – between different cities and within different public spaces in the specific city.

4.3 3D solid

The last but not least level of interpretation is analysis of quantitative parameters of 3D solid created as a 3D model of space between buildings. The 3D VOID is also resultant from geometry of forms being limitation of the shape, so geometric complexity level of triangular solid is possible to explore using the count of polygons in specific areas. The other aspect of the 3D geometry to be put under investigation is comparative analysis of volumes in certain areas of city. This responds also to parameters of capacity of particular public spaces. The 3D VOID brings also possibility of generating functional patterns of public spaces and patterns of accessibility for users getting data on functions and entrances directly from adjacent buildings. The other, more advanced scheme of interpretation, taking into consideration heights of polygons (Z coordinates) and its angle against sun direction, is identification of mostly shaded areas of public space in city. Figure 25 presents multi-polygonal structure of 3D VOID used for the introduced level of interpretation.

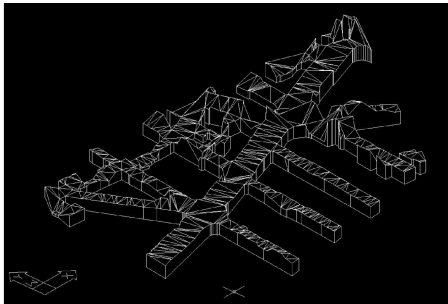


Figure 25: 3D VOID level of interpretation of complexity and spatial parameters of space between buildings.
Finally the 3D VOID can be applied as a geometric boundary for application in simulative

environment of Agent-Based Modelling of pedestrian flow etc.

Table 1: 3D VOID interpretation matrix for geometry and complexity of space between buildings.

Level of interpretation	Parametric data
Point cloud	<ul style="list-style-type: none"> - concentration of points: open space to built-up area ratio; - functional pattern of points: functional attractiveness / accessibility for users; - spectrum of points: geometrical regularities of public spaces in city-wide extent; - X-Ray observation: geometrical flow of points at bottom / top layers of point cloud
Cut surface image	<ul style="list-style-type: none"> - linear connections/graphs: regularity and typology of public spaces in comparison to other cities; - combination of linear schemes and solid objects: proportions and geometric shape of public spaces;
3D solid / 3D VOID	<ul style="list-style-type: none"> -polygon count: complexity level; -adjacent faces data: functional patterns of public spaces in city; -Z coord. and angle of polygons: shading parameters of public spaces

5. CONCLUSIONS

The presented analytic approach for physical parameters of space between buildings, called 3D VOID, brought into environment of 3D virtual city models, contributes to issues of importance of public spaces in urban areas, general trend of parametric computations, rapid development of computer technologies and general analysis of cityscape. The research initiates deeper exploration of unexplored city database hidden in geometry of its in-between spaces. The concept aims to introduce multi-aspect computer-based analytic tool into the environment of standardized and optimized 3D

virtual city models, particularly to its still insufficiently explored part of space between buildings. Geometric interpretation of this space for purpose of advanced parametric urban analyses is directed towards anticipation of extended use of virtual models technology for sustainable development of cities.

ACKNOWLEDGMENTS

This research was funded by a Norwegian Financing Mechanism. Digital model of Berlin provided by Berlin Partner GmbH. All other 3D city models used for simulation were made by Cyber Urban Center at WPUT Szczecin. I gratefully acknowledge this support.

REFERENCES

- [1] P. M. Carranza, D. Koch. A computational method for generating convex maps using the medial axis transform. In *Proceedings of 9th International Space Syntax Symposium*, Paper no. 064, Sejong University, Seoul, 2013
- [2] K. Czyńska. Attractiveness of urban landscape versus tall buildings – on the example of selected European cities. *Space&FORM* no. 22, 2014, Szczecin
- [3] J. Gehl. Space between buildings. using public space. The Danish Architectural Press, Copenhagen, 2003
- [4] B. Hillier, J. Hanson. The social logic of space. Cambridge University Press, Cambridge, 1984
- [5] B. Hillier, J. Hanson, J. Peponis. The syntactic analysis of settlements. *Architecture et Comportement Architecture and Behavior*, vol. 3, no. 3, 1987
- [6] B. Jiang, Ch. Claramunt, B. Karquist. An integration of space syntax into GIS for modelling urban spaces, *JAG*, volume 2, Issue 3/4, 2000, pages 161-171.
- [7] Y. O. Kim. The role of spatial configuration in spatial cognition. In *Proceedings of 3rd International Space Syntax Symposium*, Atlanta, 2001
- [8] T. Kolbe, S. Bacharach. CityGML: An Open standard for 3D city models, In: *Directions magazine*. Online edition: www.directionsmagazine.com, 2006
- [9] T. Kolbe, G. Groeger, L. Pluemer. CityGML - Interoperable access to 3D city Models. In *Proceedings of the International Symposium on Geo-information for Disaster Management*, Springer Verlag, Delft, 2005
- [10] K. Lynch. The image of the city. MIT Press, Massachusetts, 1960
- [11] A. Penn. Space Syntax and spatial cognition. Or, why the axial line? In *Proceedings of 3rd International Space Syntax Symposium*, Paper no. 11, Atlanta, 2001
- [12] M. Serra, J. Gill, P. Pinho. Unsupervised classification of evolving metropolitan street patterns. In *Proceedings of 9th International Space Syntax Symposium*, Paper no. 046, Sejong University, Seoul, 2013
- [13] F. P. Preparata and M. I. Shamos. Computational geometry: An introduction. In *Texts and Monographs in Computer Science*, Springer-Verlag, New York, 1985
- [14] P. Rubinowicz. Cyber Urban Design. *Archivolta*, vol. 3 (59) / 2013, Archivolta Publishing House Michał Stępień, pages 58-65
- [15] M. de Berg, M. Kowaluk. Triangulacja Dalaunaya. In *Geometria obliczeniowa*, WNT, Warszawa, 2007

ABOUT THE AUTHOR

1. Adam Zwoliński, PhD arch., MSc in Urban Housing Management, Assistant professor at West Pomeranian University of Technology (WPUT) Szczecin, Poland. Institute of Architecture and Spatial Planning (IAiPP). Since 2013, member of Cyber Urban Center (CUC) research team. Scientific interest in 3D city models, public spaces, movement patterns, parametric interpretation and analysis of urban space.

e-mail: azwolinski@zut.edu.pl

COMPOSITE CONCAVE CUPOLAE AS GEOMETRIC AND ARCHITECTURAL FORMS

Slobodan MIŠIĆ, Marija Đ. OBRADOVIĆ and Gordana ĐUKANOVIĆ
University of Belgrade, Serbia

ABSTRACT: In this paper, the geometry of concave cupolae has served as the starting point for generation of composite polyhedral structures, usable as formative patterns for architectural purposes. Linking paper folding geometry with geometry of polyhedra, concave cupolae are polyhedra that follow the method of generating cupolae (Johnson's solids: J3, J4 and J5), excluding convexity criterion, and omitting squares in the lateral surface. Instead of alternating triangles and squares, there appear two or more paired series of equilateral triangles. The criterion of face regularity is respected, as well as the criterion of multiple axial symmetry. Distribution of the triangles is based on strictly determined and mathematically defined parameters, which allows the creation of such structures in a way that qualifies them as an autonomous group of polyhedra – concave cupolae of II, IV, VI... (2N) sort. If we observe these structures as polyhedral surfaces (not as solids) connecting the concept of the cupola (dome) in the architectural sense of the term, with the geometrical meaning of (concave) cupola, the faces of the base polygons may be excluded. Thus we get a deltahedral structure – a shell made entirely of equilateral triangles, which is advantageous for the purpose of prefabrication. Due to the congruence of the major 2n-sided bases of concave cupolae of II sort with the minor bases of the corresponding concave cupolae of IV sort, joining these polyhedra into composite polyhedra is possible, as well as their elongation with concave antiprisms of II sort, or augmentation with concave pyramids of II sort. Based on the foregoing, we examine the possibilities of combining the considered polyhedra into unified composite structures. The obtained complex concave polyhedral shapes may be used as a structural skeleton in architectural design, and further applied as roof structures or stand-alone architectural formations.

Keywords: Concave polyhedra, cupola, deltahedron, equilateral triangle.

1. INTRODUCTION

The formation of composite polyhedral structures based on geometry of concave polyhedra of the second (and higher) sort, by joining the initial polyhedra, matching and combining their forms in order to obtain a new matrix applicable in architectural design is the main subject of this paper. The definitions of composite polyhedra are somewhat different, depending on whether the issue is accessed from mathematical or engineering point of view, so Timofeenko in [17] provides the following definition of composite polyhedra:

If a convex polyhedron with regular faces can be divided by some plane into two polyhedra

with regular faces, then it is said to be composite,

which applies only to the convex polyhedra. Considering that in this research the forms are explored in service of engineering solutions, we accepted a broader definition of composite solids, which was adopted by P. Huybers [2] and D. G. Emmerich [1], who granted composite polyhedra to be also polyhedra obtained by augmentations of the uniform polyhedra using Johnson's cupolae and rotundae [5] whereat the concave structures arise.

Composite polyhedra that are the subject of this study, occur by augmentation of concave cupolae of IV sort using concave cupolae of II sort and concave pyramids of II sort, and by

(gyro)elongations using antiprisms and concave antiprisms of II sort, i.e. using the polyhedra which belong to the family of concave polyhedra whose geometry is based on folding the deltahedral lateral surface.

2. CONCAVE M-m POLYHEDRA

The family of concave polyhedra obtained by the method based on folding the lateral surface plane net composed of double (or even number) series of equilateral triangles, consists of: concave cupolae of II sort [11], concave pyramids of II sort [11], [16], concave antiprisms of II sort [9], [15], as well as concave cupolae of IV and higher sorts [8]. The process of generating these polyhedra is similar to the formation of polygrammatic antiprisms [3] or pseudo-cylinders [7]. It is also akin to the origami technique, since the folding of planar net by the assigned edges produces 3D structures. All of them have in common that they are formed over a regular polygonal base, they are also regular faced, have multiple axial symmetry, and are obtained by polar array of unit concave cells consisting of equilateral triangles (spatial hexahedral or pentahedral cells) arranged around the axis of the solid.

Over the same polygonal base, by polar arrangement of a unit cell, two types of deltahedral belt can be formed: with the greater height (M) and with the lesser height (m). For this reason, this family of concave polyhedra is abbreviated to M-m polyhedra. Their parameters, linear and angular, are strictly determined and geometrically defined, as shown in previous studies [9], [11]. In this paper we give just a brief overview of these solids.

2.1 Concave cupolae of II sort

Concave cupolae of II sort (CC II) [11] are polyhedra that follow the method of generating cupolae (Johnson's solids: J3, J4 and J5), excluding convexity criterion, and omitting squares in the lateral surface. Instead of alternating triangles and squares, there appear two series of equilateral triangles. Such polyhedra are formed by connecting two regular polygons, n-gon and 2n-gon (adopted bases of the solid)

in parallel planes, by two rows of equilateral triangles which make lateral surface of a solid. The sort of cupola is determined by the number of rows of equilateral triangles in the lateral surface (Fig. 1).

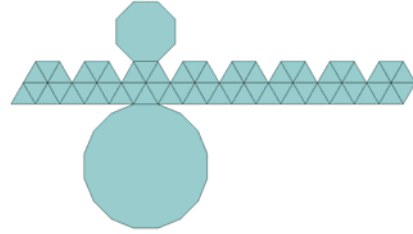


Figure 1: The lateral surface net of CC II-8

For each basic polygon, the lateral surface net can be folded in two ways: with central vertex of the unite hexahedral cell retracted, giving the major height of the solid and the first type, CC II-M, or extruded, giving the minor height of the solid, and the second type, CC II -m. There are fourteen possible representatives of CC II altogether [11], from $n=4$ to $n=10$, each of them with two types, M and m (Fig. 2).



Figure 2: CC II-8M and CC II-8m

2.2 Concave pyramids of II sort

Concave pyramids of second sort (CP II) are polyhedra which follow the method of generating Concave Cupolae of second sort (CC II) [16], using the same method of folding the double row of equilateral triangles plane net, as shown in Fig. 4 and Fig. 6. These polyhedra are also organized over a regular n-sided polygonal base and are regular faced, with deltahedral lateral surface consisting of equilateral triangles, as for Johnson's pyramids (J1 and J2). Unlike CC, the unit cell that forms the solid by its radial array now is a spatial pentahedral cell, having a common vertex in the apex of the solid.

There are two types of concave cupolae of II sort. The first type, CP II-A, has the number of unit cells equal to the number of the base polygon's sides, since it covers all the bases from $n=6$ to $n=9$, both odd and even. The second type, CP II-B, formed with the halved number of sides is possible only for the even bases, $n=6$, $n=8$, $n=10$.

The number of vertices (V), edges (E) and faces (F) for any CP II-A over n -sided polygonal base are calculated by formulae:

$$V = 3n + 1 \quad (1)$$

$$E = 8n \quad (2)$$

$$F = 5n + 1 \quad (3)$$

and for any CP II-B:

$$V = 2n + 1 \quad (4)$$

$$E = 5n \quad (5)$$

$$F = 3n + 1 \quad (6)$$

There are seven representatives of CP II-A: four CP II-MA with greater (major) height (indented central vertex O of the spatial hexahedral unit cell), Fig. 5a (from $n=6$ to $n=9$), and three CP II-mA (from $n=6$ to $n=8$) [16] with lesser (minor) height (protruding central vertex O of the spatial hexahedral unit cell), Fig. 5b, while there are six representatives of CP II-B (Fig. 7): the two types for the three possible bases, $n=6$, $n=8$ and $n=10$.



Figure 4: The lateral surface net of CP II-8A



Figure 5: CP II-8MA and CP II-8mA



Figure 6: The lateral surface net of CP II-8B

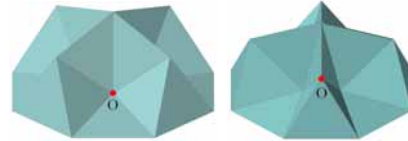


Figure 7: CP II-8MB and CP II-8mB

2.3 Concave antiprisms of II sort

Concave antiprisms of the second sort (CA II) are polyhedra consisting of two congruent regular polygonal bases in parallel planes and deltahedral lateral surface [9], [15]. The lateral surface consists of two-row strip of equilateral triangles, arranged in such a way to form spatial hexahedral elements, by which polar arrangement around the axis that connects bases' centroids, the polyhedron is created (Fig. 8-9).



Figure 8: The lateral surface net of CA II-10

Over the same polygonal base the two different types of concave antiprisms of II sort can be formed: CA II-M, with the greater height (indented central vertex O of the spatial hexahedral unit cell), and CA II-m, with lesser height (protruding central vertex O of the spatial hexahedral unit cell), Fig. 9.



Figure 9: CA II-8M and CA II-8m

The constructive procedure for generating CA II and determining their metric relations was elaborated in detail [9], [15]. The number of vertices (V), edges (E) and faces (F) for any CA II over n -sided polygonal base is calculated by formulae:

$$V = 4n \quad (7)$$

$$E = 10n \quad (8)$$

$$F = 6n + 2 \quad (9)$$

Concave antiprisms of II sort can be formed over any regular polygonal basis so they represent an infinite family of concave regular faced polyhedra.

2.4 Concave cupolae of IV sort

Concave Cupolae of IV sort (CC IV) are polyhedra which follow the method of generating CC II, only with four-fold strip of equilateral triangles in the plane net, and with the additional quadrilateral cells inserted between the hexahedral cells in the middle zone of the lateral surface, in order to enclose the inner space, and to form a deltahedral shell modeled by strictly geometrically and mathematically defined principles [8]. CC IV is formed, like CC II, over the regular polygonal bases, n -sided (Ω_1) and $2n$ -sided (Ω_2), which lie in parallel planes (Fig. 10).

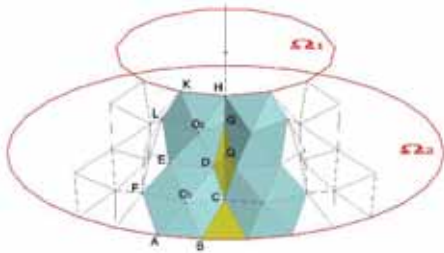


Figure 10: The generation of CC IV by the polar array of the spatial hexahedral unit cells

The positions and the heights of the vertices of CC IV are constructed using the intersections of the vertical plane β (Fig. 11), determined by the axis of the solid and the vertex of the base polygon with the spheres of radii $r=a$

(equal to the edge length of the solid) Centers of the spheres are set in the neighboring vertices of the spatial hexahedral cell (Fig. 11). For more detailed description of the constructive procedure see in [9].

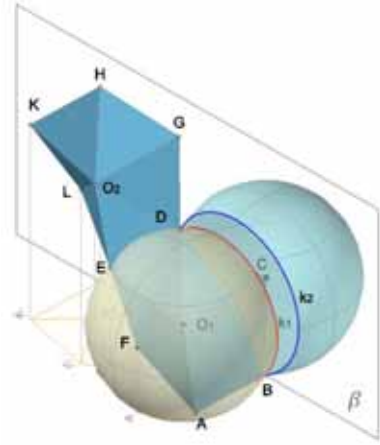


Figure 11: 3D model of the CC IV unit cell, the construction of the vertices B and D positions for the spatial hexahedral cell ABCDEFO₁ [9]

In the example of four different types of CC IV with the same polygonal base, shown in Figure 12, we can notice that their fundamental difference lies in the positions of the central vertices (O₁ and O₂) of the spatial hexahedral unit cell, which may be indented or protruding, as seen from the outside.

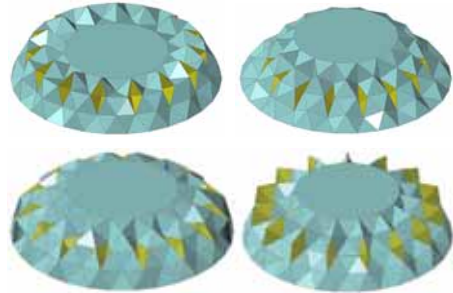


Figure 12: a) CC IV-15Mm, b) CC IV-15mm, c) CC IV-15mM, d) CC IV-15MM [9]

2.5 Concave Cupolae of the higher sorts

The sort of the cupola is determined by the number of rows of equilateral triangles in the lateral surface forming the spatial unit cells by arrangement around a common vertex, which requires two linked series of triangles. Consequently, the concave cupolae are always of the even sort [9].

The number of vertices (V), edges (E) and faces (F) for any CC of the sort x over n -sided polygonal base is calculated by formulae [9]:

$$V = \frac{5x}{2} \cdot n \quad (10)$$

$$E = \left(\frac{15x}{2} - 3 \right) \cdot n \quad (11)$$

$$F = (5x - 3) \cdot n + 2 \quad (12)$$

For each sort of concave cupolae we can form the lateral shell (with the even number of equilateral triangles' rows) over eleven additional polygonal bases, in relation to the CC of the lower sort [9]. The only exception are CC II, because they can be generated over seven different polygonal bases ($4 \leq n \leq 10$), as described in [11]. Also, CC II are an exception for the fact that they are the only concave cupolae that do not have bonding spatial quadrihedral cells in their lateral surface net.

The number (x) of the possible different concave cupolae of the sort v over the same polygonal base is calculated according to the formula:

$$x = 2^{\frac{v}{2}} \quad (13)$$

3. COMPOSITE POLYHEDRAL STRUCTURES BASED ON GEOMETRY OF CONCAVE M-m POLYHEDRA

3.1 Augmentations

Augmented polyhedron is a polyhedron with one or more other solids adjoined [4]. At the regular faces of polyhedron, we can subjoin other figures with identical bases to the respective face of the polyhedron. Most often the augmentation involves adding pyramids (regular-faced, though not necessarily) (P1, J1, J2),

cupolae (J3, J4, J5) or pentagonal rotunda (J6) [5]. In this paper, we used CC IV as the initial figure which can be augmented by CC II and by CP II.

Table 1: Possible mutual combination of CC II

CC		II-4	II-5
		m M m M	
II-4			
II-5			
II-6			
II-7			
II-8	m M	• •	
II-9			
II-10	m M		• •

Table 2: Possible combinations of CC II and CC IV

CC		II-6	II-7	II-8	II-9	II-10
		m M m M m M m M m M				
IV-11						
IV-12	Mm	x x				
	mm	• •				
	mM					
	MM	• •				
IV-13						
IV-14	Mm		• •			
	mm		• •			
	mM		• •			
	MM		• •			
IV-15						
IV-16	Mm			• •		
	mm			x x		
	mM			• •		
	MM			• •		
IV-17						
IV-18	Mm				x x	
	mm				x x	
	mM			• •		
	MM			• •		
IV-19						
IV-20	Mm					x x
	mm					
	mM					• •
	MM					
IV-21						

The research results on mutual combination of CC II are shown in Table 1, while the possible combinations of CC II and CC IV are shown in

Table 2, providing the 32 possible cases of matching pairs of concave cupolae.

Matching the congruent polygonal faces (bases) of CC IV, CC II and CP II, we obtain polyhedral structures with compositive deltahedral shell, which completely enclose the space confined by the polygonal basis.

The composite polyhedra formed in this manner, which could be (according to Gestalt principles) related with the form cupola, i.e. dome in the architectural sense, in Tab. 1-2 are marked with (•). The cases of just deltahedral lateral surface (not the whole solid) are marked with (x). Thus formed structure can be further augmented by CP II-A. In the case of CC II basic n -sided polygon being $n=6$, $n=8$ or $n=10$, we may also augment the observed cupola by CP II-B. The representatives of CC IV cannot be cross-combined because in the span of $11 \leq n \leq 21$ there are no representatives A and B such that $2n_A \equiv n_B$. The composite polyhedral structure consisting of CC IV-16mm, CC II-8m and CP II-8B, is presented in Fig. 13-14.

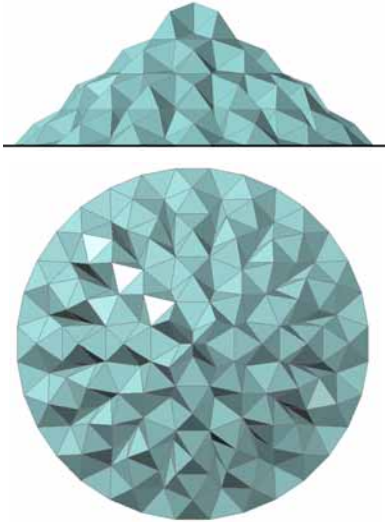


Figure 13: Front view and top view of the composite polyhedron obtained by joining CC IV-16mm, CC II-8m and KP II-8B

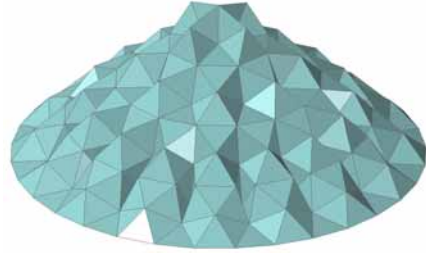


Figure 14: 3D model of composite polyhedron obtained by joining CC IV-16mm, CC II-8m and CP II-8B

This paper illustrates just a few examples of composite structures based on geometry of concave M-m polyhedra, having a form associative with the domes (cupolas) in the architectural sense. A complete gallery would require far more space, since all variations, including augmentation and (gyro)elongations, far exceed the 32 cases given in Tables 1 and 2.

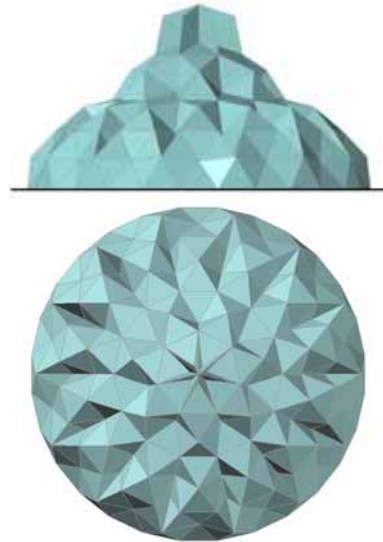


Figure 15: Front view and top view of the composite polyhedron obtained by joining CC IV-14mM, CC II-7m and CP II-7MA



Figure 16: 3D model of composite polyhedron obtained by joining CC IV-14mM, CC II-7m and CP II-7MA



Figure 17: Front and top view of composite polyhedron obtained by joining CC IV-14MM, CC II-7M i CP II-7mA

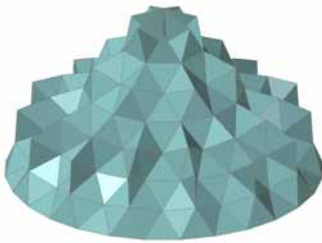


Figure 18: 3D model of composite polyhedron obtained by joining CC IV-14MM, CC II-7M and CP II-7mA



Figure 19: Front and top view of composite polyhedron obtained by joining CC IV-16Mm, AP-16, CC II-8M and CP II-8mA

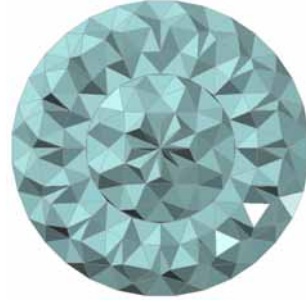


Figure 20: 3D model of composite polyhedron obtained by joining CC IV-16Mm, AP-16, CC II-8M and CP II-8mA

3.2 Elongations

In Johnson's classification of convex polyhedra [5] there are 35 solids obtained by elongations or gyro-elongations (bi)pyramids J1 and J2 or (bi)cupolae J3, J4 and J5.

Elongation is a method of joining prisms of the congruent bases to the base of the initial solid, while gyro-elongation is the method of joining antiprisms.

In this paper, we expanded the meaning of elongation by including CA II, with bases congruent to the bases of the initial solids. Since the sides of CA II bases are identical and mutually parallel, we have the cases of (conca) elongations [13], and not gyro-elongations.

Using the method of elongation, the solids can be additionally extended, resulting in an increase of the surface area / volume of the new structures, in order to provide larger interior space, or contribute to harmonious proportions, for the aesthetic reasons. In order to achieve the desired goal, all lateral surfaces transformed into deltahedral, in this paper we consider only the cases of elongation by surfaces composed strictly of equilateral triangles:

- a) antiprisms (Fig. 21-22)
- b) CA II (Fig. 23-24)

A group of polyhedra obtained by elongations and gyro-elongations of CC II is described in [9], [13], while a group of polyhedra obtained by elongations and gyro-elongations of CP II is described in [16]. For elongation of CC IV, CC II and CP II we use deltahedral belts formed over regular polygons congruent with the bases of CC IV which is (gyro)elongated.

The composite polyhedron obtained by joining CC IV-16mM, AP-16, CC II-8m and CP-8B is presented in Fig. 21-22.

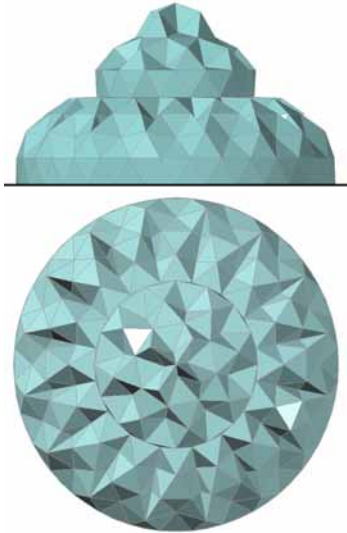


Figure 21: Front and top view of the composite polyhedron obtained by joining AP-32, CC IV-16mM, AP-16, CC II-8m and CP-8B

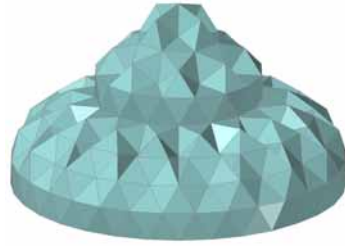


Figure 22: 3D model of composite polyhedron obtained by joining AP-32, CC IV-16mM, AP-16, CC II-8m and CP-8B

As noted above, over the same polygonal base it is always possible to construct two different types of concave antiprisms of second sort: CA II-M and CA-II m. Therefore, the observed composite cupola, including antiprisms, can be elongated in three different ways, so the number of composite polyhedral structures formed increases as a multiplying function of three for each polyhedron participating in the composition.

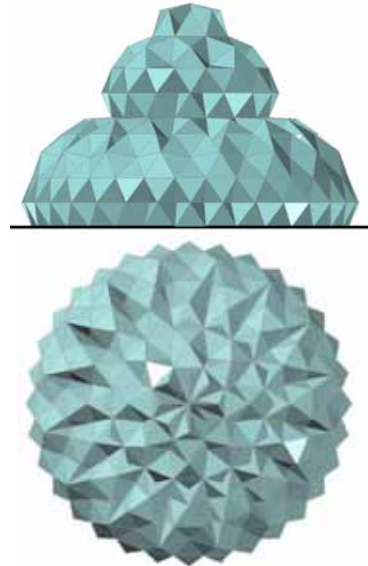


Figure 23: Front and top view of the composite polyhedron obtained by joining CA II -32m, CC IV-16mM, CA-16m, CC II-8m, CP-8mA



Figure 24: 3D model of composite polyhedron obtained by joining CA II -32m, CC IV-16mM, CA-16m, CC II-8m and CP-8mA

4. APPLICABILITY OF COMPOSITE M-m STRUCTURES IN ARCHITECTURAL PRACTICE

The research on possibilities of concave cupola's application in architectural practice [9], [11], [14], has revealed a number of advantages. Primarily, there is a deltahedral surface, wherein the presence of equilateral triangles gives the construction the necessary unification, which behooves prefabrication. Due to the congruence i.e. commonality of lateral faces in the surface, the serial production of prefabricated elements is possible, which could easily form a structure that follows the geometry of these polyhedra.

The deltahedral shell itself can be formed by mounting triangular panels, frames, reinforced slabs, precast, y-profile rails, or even spatial tetrahedral grids, which allow spanning much larger range. Also, the triangle as a stable geometric figure enables organization of the structure into geometric form with the exact position of vertices, which further ensures immobility of the construction. Based on the analysis in [14] it was shown that the structure based on the geometry of the concave cupolae has good dynamic and static properties. Besides, the possible use of diagrid structural system enables the free organization of interior space.

One of the deficiencies of concave cupolae as engineering forms is the existence of large flat surfaces of polygonal bases. This was the primary motivation of our research of new

composite polyhedral structures completely converted into deltahedral lateral surfaces, whose geometry would come closer to the form of architectural dome. Herewith, the architectural potential of polyhedral structures [6] is once again confirmed.

5. CONCLUSIONS

The concave M-m polyhedra, including CC II, CP II, CA II and CC IV, having common polygonal bases, can be joined, merged, combined and integrated into a whole, and therefore are suitable for the formation of composite polyhedral structures. In this paper, we have explored possible solutions whose shape would be associative to the dome in the architectural sense. For this purpose, we started from CC IV, which we transformed by the procedures of augmentation and (gyro)elongations into complex polyhedral forms. By the augmentations of CC IV and CC II we obtained deltahedral surface which completely encloses interior space. A possible need for greater height or volume can be successfully met by elongation of the resulting structure either by antiprisms or by CA II. In addition, all the desirable geometrical, static and aesthetic traits of concave cupola (as an architectural form) are transferred to the newly created composite polyhedral structure.

ACKNOWLEDGMENTS

The research is supported by the Ministry of Science and Education of the Republic of Serbia, Grant No. III 44006.

REFERENCES

- [1] D. G. Emmerich. Composite polyhedral. *Int. Journal of Space Structures*, 5, pages 281-296. 1990.
- [2] P. Huybers. Form generation of polyhedral building shapes. *Int. Journal of Space Structures, Special issue on Morphology and Architecture, Vol. 11, No 1&2*, pages 173-181. 1996.
- [3] P. Huybers. Prism Based Structural forms, *Engineering Structures* 23, pages 12-21. Elsevier, 2001

- [4] P. Huybers. Polyhedroids. *An Anthology of Structural Morphology* edited by René Motro, pages 49-62. World Scientific Publishing Co. Pte. Ltd., 2009.
- [5] N.W. Johnson. Convex Solids with Regular Faces, *Canadian Journal of Mathematics*, Vol 18, pages 169-200, University of Toronto Press, Toronto, Canada, 1966.
- [6] I. J. Markov and J. F. Gabriel. Spatial and structural aspects of polyhedra, *Engineering Structures* 23, pages 4-11, Elsevier, 2001
- [7] K. Miura. Proposition of Pseudo - Cylindrical Concave Polyhedral Shells, Institute of Space and Aeronautical Science, University of Tokyo, Report No.442, Tokyo, 1969.
- [8] S. Mišić and M. Obradović. Forming the cupolae with concave polyhedral surfaces by corrugating a fourfold strip of equilateral triangles. *Proceedings of 2nd International Scientific Conference MoNGeometrija, 2010.*, pg. 363-374, Beograd, Serbia, 2010.
- [9] S. Mišić. Constructive – geometric generating of cupolae with concave polyhedral surfaces. Doctoral Disertation, University of Belgrade – Faculty of Architecture, 2012.
- [10] S. Mišić, M. Obradović and B. Popkonstantinović. The structural transformation of concave cupolae of fourth sort using different variants of constructive procedure. *Proceedings of 4th International Scientific Conference Mongeometrija 2014.* pp 148-158, Vlasina, Serbia, 2014.
- [11] M. Obradović. Constructive - geometrial elaboration on toroidal deltahedra with regular polygonal bases, Doctoral Disertation, University of Belgrade – Faculty of Architecture, 2006.
- [12] M. Obradović and S. Mišić. Concave Regular Faced Cupolae of Second Sort. 13th ICGG, *Proceedings of International Conference on Geometry and Graphic, on CD*, Dresden, Germany, July 2008.
- [13] M. Obradović. A group of polyhedra arised as variations of concave bicutolae of second sort. *Proceedings of 3rd International Scientific Conference Mongeometrija 2012*, pages 95-108. Novi Sad, 2012.
- [14] M. Obradović, S. Mišić, B. Popkonstantinović, M. Petrović, B. Malešević, R. Obradović. Investigation of Concave Cupolae Based Polyhedral Structures and Their Potential Application in Architecture, *TTEM, Vol.8., No.3*, pages 1198-1214, DRUNPP, Sarajevo, 2013.
- [15] M. Obradović, B. Popkonstantinović and S. Mišić. On the Properties of the Concave Antiprisms of Second Sort. *FME Transactions Vol. 41, No 3*, pages 256-263, Faculty of Mechanical Engineering, Belgrade, 2013
- [16] M. Obradović, S. Mišić and B. Popkonstantinović. Concave pyramids of second sort – the occurrence, types, variations. *Proceedings of 4th International Scientific Conference Mongeometrija 2014.* pp 159-170, Vlasina, Serbia, 2014.
- [17] A. V. Timofeenko. The non-Platonic and non-Archimedean noncomposite polyhedral. *Journal of Mathematical Sciences, Vol. 162, No 5*, pages 710-729, Springer Science+Business Media, Inc., 2009.

ABOUT THE AUTHORS

1. **Slobodan Mišić** PhD, Assistant professor at the University of Belgrade - Faculty Civil Engineering. In his scientific reserch, he is dealing mostly with constructive geometry and polyhedral structures in architecture and engineering. slobodan@grf.bg.ac.rs
2. **Marija Đ. Obradović** PhD, Associate professor at the University of Belgrade - Faculty Civil Engineering, teaching Computational Geometry and Descriptive Geometry. Her scientific interest is related mainly to constructive geometry and polyhedral structures in architecture and engineering. marijao@grf.bg.ac.rs
3. **Gordana Đukanović** PhD, Assistant professor at the University of Belgrade – Faculty of Forestry. In her scientific reserch, she is dealing with descriptive and relativistic geometry. gorddana@rcub.bg.ac.rs

Computing the Higher Dimensional Delaunay Decomposition using Depth-First Search

Martin HÜNNIGER

Technical University Zwickau, Germany

ABSTRACT: This paper presents a new $O(n^2)$ -time algorithm for the computation of the Delaunay decomposition in \mathbb{R}^d . The algorithm does so by exploring the d -faces of the Delaunay decomposition only. It combines Gärtner's algorithm for computing the smallest enclosing ball of a point set with a complete search strategy such as DFS or BFS.

Keywords: Geometry, Delaunay decomposition, multidimensional, graphics

1. INTRODUCTION

The Delaunay decomposition of a point set with n elements of \mathbb{R}^d is a common and well understood problem in computational geometry. The data structure is at the core of many other geometrical algorithms such as surface/manifold reconstruction, point location or in the use of finite element methods. The Delaunay decomposition is an optimal data structure in terms of the shape of its faces and it is dual to the Voronoi decomposition. The latter is another important data structure used in different computational disciplines ranging from Biology and Chemistry to Astronomy and Architecture.

For low dimensions many efficient implementations of algorithms to compute the Delaunay decomposition exist. In two dimensions the fastest implementation uses $O(n \log n)$ time with $O(n)$ space. This can be achieved by the flip algorithm [2] or by an incremental construction of the Delaunay triangulation [3]. In three dimensions, there are also well studied algorithms, extending the flip paradigm [4] or the incremental construction [5].

Known algorithms for the computation of the Delaunay decomposition in d dimensions exist and are featured in software packages such as QHULL [6] or the CGAL library [7]. Both use a fast $(d+1)$ -dimensional convex hull algo-

rithm to compute the Delaunay decomposition by lifting the d -dimensional Points on a $(d+1)$ -paraboloid. Both have running times complexity of $O(n \log n + n^{\lceil d/2 \rceil})$

Any Delaunay decomposition of n points in \mathbb{R}^d contains $O(n^{\lceil d/2 \rceil})$ simplices, which becomes a computational problem if the dimension gets higher. To avoid this, a new technique is introduced: we only compute the $O(n)$ d -dimensional Delaunay simplices and store them in a data structure that allows efficient access to their vertices. Computing their intersections generates the lower dimensional faces. We thus circumvent the $O(n^{\lceil d/2 \rceil})$ term dominating the running time of the other algorithms and achieve faster results.

The algorithm we present is also interesting in its own right, since it connects the geometrical problem of finding d -simplices with the combinatorial approach of depth-first search even with the option of parallelizability. Since we only compute the Delaunay d -simplices, we achieve a running time of $O(n^2)$, which is comparable with time complexities of the algorithms mentioned before.

This paper is organized as follows: In section 2 we feature basic definitions and facts about the Delaunay decomposition and Voronoi Diagrams as well as techniques used in later parts of the

text. Section 3 gives an introduction to the algorithm. The correctness and running time analysis is provided in section 4 and a brief discussion follows in section 5.

2. FUNDAMENTALS

Throughout this paper let $P = \{p_0, \dots, p_n\}$ be a subset of \mathbb{R}^d . A subset $\sigma = \{x_0, \dots, x_k\}$ of P is called *Delaunay k -simplex*, $0 \leq k \leq d$, if the circumball of σ contains no other points of P than x_0, \dots, x_k . This property is called the *empty ball property*. The collection of all Delaunay d -simplices and their faces is called *Delaunay decomposition*, $\mathcal{D}(P)$.

Please observe that any simplex of $\mathcal{D}(P)$ is defined only by its vertices. A simplex is a face of a second simplex iff it is a subset, and a simplex is a common face of two simplices iff it is contained in their intersection. In this way incidences between simplices may be computed by only accessing their common vertices.

For each k -simplex $\sigma = (x_0, x_1, \dots, x_k)$ we may compute geometrical properties by matrix A_σ that we define as

$$A_\sigma = \begin{pmatrix} x_{11} - x_{01} & \dots & x_{k1} - x_{01} \\ x_{12} - x_{02} & \dots & x_{k2} - x_{02} \\ \vdots & \ddots & \vdots \\ x_{1d} - x_{0d} & \dots & x_{kd} - x_{0d} \end{pmatrix} \quad (1)$$

Another geometrical decomposition of utmost importance is the Voronoi diagram $\mathcal{V}(P)$ of the point set $P = \{p_0, \dots, p_{m-1}\} \subseteq \mathbb{R}^n$. It is composed of so called *Voronoi cells*, where each Voronoi cell V_q of a point $q \in P$ is defined as

$$V_q = \{x \in \mathbb{R}^d : \|x - q\| \leq \|x - p\|, p \in P\}$$

The Voronoi cell of a set $Q = \{q_0, \dots, q_k\} \subseteq P$ of $k + 1$ points is then defined as

$$V_Q = \bigcap_{q \in Q} V_q$$

We point out the well known duality between Delaunay decomposition and Voronoi diagram by the transformation

$$\sigma = \{x_0, \dots, x_k\} \in \mathcal{D}(P) \Leftrightarrow V_{\{x_0, \dots, x_k\}} \in \mathcal{V}(P)$$

Our algorithm computes and stores the Delaunay d -simplices explicitly. We maintain the empty ball property for each simplex in a way that is described by Bernd Gärtner in [1]. Let x_0, \dots, x_k be the vertices of an arbitrary simplex σ for $0 \leq k \leq d$. We store the vectors $a_i = x_i - x_0$ in the matrix $A_\sigma = (a_1, \dots, a_k)$ as defined in equation 1. The system of equations given by this matrix may be used to compute the circumball of σ and its circumcenter c_σ , as well as the affine coordinates of c_σ with respect to a_1, \dots, a_k . It is possible to accelerate all computations connected to A_σ by maintaining it as a *QR* decomposition. We explain this fact in more detail in section 4.2. The reader may also refer to Gärtner's article for a full discussion on the topic.

The algorithm to be described in this article aims to generate Delaunay simplices. We use the technique above to maintain the empty ball property for all simplices we generate.

Let us assume we already found a Delaunay d -simplex σ . Therefore, its circumball has the empty ball property and we may use A_σ to compute its center c_σ . By dropping an arbitrary vertex x from σ , a simplex σ' is formed. Conforming again to Gärtner, we may *incrementally* compute $A_{\sigma'}$ and use it to deduce the affine space spanned by σ' . Projecting c_σ onto the space spanned by σ' defines a point p and a line ℓ . This line contains the Voronoi edge given by the vertices of σ' , since these vertices are equidistant to all points in ℓ . Let $\delta = p - c_\sigma$. Moving the center of the circumball $B_{\sigma'}$ in the direction of δ , *away* from the removed point x along ℓ , grows the ball until a new point x' of P enters its boundary. If σ' is a simplex of the border of $\mathcal{D}(P)$ we modify the algorithm to include an imaginary point at infinity for consistency.

We need to determine a point x' when forming a new Delaunay d -simplex. This is done as follows: For each point $y \in P$ we may compute the value t_y that gives the "time" when y enters $B_{\sigma'}$ as we move its center along ℓ . Let q be any vertex of σ' as they are always in $B_{\sigma'}$. The ball

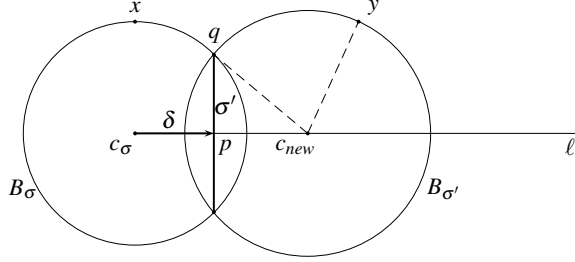


Figure 1: The situation when moving the circumcircle along line ℓ . The point y enters $B_{\sigma'}$.

around c_σ when y enters it satisfies

$$\|q - (c_\sigma + t_y \delta)\|^2 = \|y - (c_\sigma + t_y \delta)\|^2$$

which is equivalent to

$$\|q\|^2 - 2q^T(c_\sigma + t_y \delta) = \|y\|^2 - 2y^T(c_\sigma + t_y \delta)$$

and finally leads to

$$t_y = \frac{\|y\|^2 - \|q\|^2 - 2c_{\sigma'}^T(y - q)}{2\delta^T(y - q)} \quad (2)$$

We may then choose the point y with smallest t_y as the point that enters $B_{\sigma'}$ first, thus determining a new Delaunay d -simplex τ adjacent to σ via the common face σ' . This proves:

Lemma 1 *Let σ be Delaunay d -simplex defined by P and let x be a vertex of σ . Define the circumcenter of σ as c_σ and δ the perpendicular of c_σ to the affine hull of $\sigma \setminus \{x\}$. Then the adjacent Delaunay d -simplex to σ is given by $\sigma' = (\sigma \setminus \{x\}) \cup \{x'\}$, where*

$$x' = \arg \min_{y \in P \setminus \sigma} \{t_y\}$$

How do we find an initial Delaunay d -simplex? We describe an iterative procedure that maintains a Delaunay k -simplex for $0 \leq k \leq n$ and extends it by adding points from P to it. To start, we choose a point from inside the convex hull $CH(P)$ of P , which is not in P . Call this

point c . An initial Delaunay 0-simplex σ is the point $x \in P$ closest to c . An invariant of this procedure is that the simplex σ always spans an affine subspace of \mathbb{R}^d onto which we project c thus getting a point p and a line ℓ . Unlike in the procedure of finding adjacent simplices, we move c along ℓ in direction of $-\delta = c - p$. Thereby, we enlarge the ball B_σ determined by c and the vertices of σ . B_σ eventually reaches a new point y that we insert into σ . We then repeat the process.

This process can not be continued if c lies inside the affine space spanned by σ . If this is the case and σ is d -dimensional, then we found a Delaunay d -simplex. Otherwise, if σ is not d -dimensional, we either perturb the position of c in an ε -neighbourhood or we could restart the process with an other choice for x . However, as the probability for such a situation is 0, it is not significant to the running time or correctness of the algorithm. Also observe, that we will never run off to infinity as we move along ℓ , since we chose c to be inside the convex hull of P .

Therefore we may say:

Lemma 2 *Let c be a point inside the convex hull $CH(P)$. If c lies inside a Delaunay d -simplex σ , then σ can be found as an initial Delaunay simplex.* \square

3. THE ALGORITHM

The algorithm works in two phases. At first it tries to find an initial Delaunay simplex and

when it is found, it enters a search stage that explores the whole Delaunay decomposition expanding Delaunay d -simplices as described in the last section.

3.1 The Initial Phase

Given a finite set of points $P \subseteq \mathbb{R}^d$, the algorithm tries to determine a d -simplex of $\mathcal{D}(P)$.

1. Choose a point c inside $CH(P) \setminus P$.
2. Find the closest point $x \in P$ to c .
3. Let $p = x$, $\sigma = (x)$, and let ℓ be the line through p and c .
4. While σ does not contain $d+1$ vertices do
 - (a) Define the point p by projecting the center c_σ onto the affine space spanned by σ' .
 - (b) Define the line ℓ by p and c_σ .
 - (c) Define the direction $-\delta = c - p$
 - (d) Move the ball defined by the vertices of σ' and center $c_{\sigma'}$ along ℓ in direction $-\delta$ until the first point y enters the ball or the point at infinity is reached. (refer to equation 2).
 - (e) Insert y into σ .
 - (f) Let p the projection of c onto the affine space spanned by σ .
 - (g) Let ℓ be the line through p and c .
5. Return σ .

A convenient way to find an initial point c is to select an extreme point q_0 of P and find a point $q_1 \in P$ of maximal distance to q_0 . Let $c = \frac{1}{2}(q_0 + q_1)$, then it always lies inside the convex hull of P .

3.2 The Search Phase

Given a finite set of points $P \subseteq \mathbb{R}^d$ and a d -simplex σ of $\mathcal{D}(P)$, the following algorithm computes all Delaunay d -simplices adjacent to σ :

1. For all vertices x of σ do
 - (a) Define the $(d-1)$ -simplex σ' by removing x from σ .
 - (b) Define the point p by projecting the center c_σ onto the affine space spanned by σ' .
 - (c) Define the line ℓ by p and c_σ .
 - (d) Define the direction $\delta = p - c$
 - (e) Move the ball defined by the vertices of σ' and center $c_{\sigma'}$ along ℓ in direction δ (away from x) until the first point y enters the ball or the point at infinity is reached.
 - (f) If y is not the point at infinity then define the new d -simplex τ by inserting y into σ' and append it to the list L of d -simplices adjacent to σ .
2. Return L .

This algorithm serves as an expansion step during a search algorithm. It is possible to use any (complete) search strategy for the exploration of the space of Delaunay d -simplices. We may give each d -simplex a mark to avoid multiple exploration of the same simplex. For the same reason, we need to mark the simplices containing the point at infinity.

4. ANALYSIS

Throughout this section we assume that points are in general position. This may be achieved by perturbing the input by a small amount $\varepsilon > 0$.

4.1 Correctness

Lemma 3 *The initial phase algorithm correctly computes an initial Delaunay simplex.*

Proof. Starting out with a single point c we define a 0-simplex σ by the point $x \in P$ closest to c . Therefore, the circumball B through σ with center c contains no other points from P . It may be possible that another point x' has the same distance from c as x but in this case it gets inserted

into σ during the next iteration. In any case, B has the empty ball property.

During the iteration of the process the circum-ball B upholds the empty ball property. As long as σ has less than $d + 1$ vertices, the projection of c onto the affine hull of σ defines a line ℓ . As we expand B by moving c along ℓ , we take the first point from P that enters B as a vertex into the new simplex σ' . Therefore, the empty ball property is satisfied, proving that σ' is a Delaunay simplex.

The iteration ends with a d -dimensional Delaunay simplex. \square

Lemma 4 *Given a Delaunay simplex, the expansion algorithm correctly generate all Delaunay simplices adjacent to it.*

Proof. Given a Delaunay d -simplex σ , we maintain the empty ball property by removing a vertex x from σ and inserting the point y that firstly enters the circumball defined by σ' with center c_σ as we move its center along ℓ away from x in direction δ . If no such point exists in P , the point at infinity is used as y . In either instance the d -simplex τ defined by σ' and y is a Delaunay d -simplex. \square

Theorem 1 *A search algorithm, for instance DFS, utilizing the expansion algorithm on an initial Delaunay d -simplex computes $\mathcal{D}(P)$.*

Proof. Observe that the space we search is finite since P is finite. So, when using DFS the process will terminate as long as we do not run into undetected cycles, which can be prevented by marking already explored d -simplices.

It remains to show, that we explore all of the Delaunay d -simplices during the search. This follows from the fact that the expansion algorithm applied to a d -simplex finds all its neighbours and from the completeness of DFS. \square

We prefer to use DFS in this context because it features a better memory footprint. Provided enough memory, any complete search strategy would suffice.

4.2 Running Time

To analyze the running time, we have to take a closer look at the representation of simplices. Since the following operations are mandatory, we need to find an efficient way to carry them out:

- Insertion of a point into a simplex
- Removal of a point from a simplex
- Computation of the circumcenter
- Projection of a point onto an affine space

All these operations scale with the dimension d . As it is pointed out in Gärtner's paper [1], the running time is $O(d^2)$ for any of these primitives.

We use the matrix A_σ from equation 1 to represent the affine space spanned by the vertices of σ . The projection and the computation of the circumcenter involves solving a system of linear equations. We use a QR decomposition to solve these systems. This composition is updated on insertion and removal of points of σ using Givens Rotations. We therefore store two matrices Q and R both with d^2 entries. The vectors representing $\sigma = (x_0, \dots, x_k)$ are always of the form $x_i - x_0$ for $1 \leq i \leq k$. By this convention, we need $O(d^2)$ time for insertion and deletion.

The time needed to compute the circumcenter of σ is $O(d^2)$ as an easy computation shows. The circumcenter c of the k -simplex σ satisfies the following equations:

$$(x_i - c)^T (x_i - c) = r^2, \quad i = 0, \dots, k$$

$$\sum_{i=0}^k \lambda_i x_i = c$$

$$\sum_{i=0}^k \lambda_i = 1$$

With the definitions of $a_i = x_i - x_0$ and $C = c - x_0$, these equations may be rewritten as:

$$\begin{aligned} C^T C &= r^2 \\ (a_i - C)^T (a_i - C) &= r^2, \quad i = 0, \dots, k \quad (3) \\ \sum_{i=0}^k \lambda_i a_i &= C \end{aligned}$$

And by substituting C for $\sum_{i=0}^k \lambda_i a_i$ in 3 we get with the former definition of $A_\sigma = (a_1, \dots, a_k)$:

$$2A_\sigma A_\sigma^T \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_k \end{pmatrix} = A_\sigma^T A_\sigma \quad (4)$$

Solving equation 4, we get the affine coefficients for C . Since we maintain a QR decomposition of A_σ this operation involves $O(d^2)$ steps.

These coefficients determine, in which direction we need to expand the circumball on ℓ : If the λ_i , that corresponds to the vertex x we removed from σ , is positive then c_σ lies on the same side of the affine space as x .

The projection of any point, including the circumcenter, onto the affine space spanned by σ also involves solving the linear system represented in Q and R . Thus it can be performed in $O(d^2)$ time. It is also described by Gärtner [1], that the QR decomposition behaves nicely with respect to numerical stability using standard floating point arithmetic when using Givens rotations for the updates of the factors Q and R .

Having analyzed these primitive operations we now may turn the attention to the algorithms presented in the last section.

Lemma 5 *During the initial phase, a Delaunay d -simplex is found using $O(nd^3)$ operations.*

Proof. During the construction of the Delaunay d -simplex σ , the procedure of finding an initial point c and the first vertex x of σ involves $O(nd)$ operations. This is because we need to check distances for all points in P . Projecting the center

c onto the affine span of $\sigma \setminus \{x\}$ giving p needs $O(d^2)$ time.

The evaluation of equation (2) for all points of P is $O(nd^2)$. These costs arise during the insertion of each vertex. Since there are d more vertices to find, a whole of $O(nd^3)$ operations are needed. \square

Lemma 6 *For any Delaunay d -simplex σ , the expansion algorithm takes $O(nd^3)$ operations to compute all Delaunay d -simplices adjacent to σ .*

Proof. The evaluation of equation (2) takes $O(nd^2)$ operations and is performed for any vertex of σ that gets dropped. There are $d + 1$ vertices and thus $O(nd^3)$ total operations. \square

Theorem 2 *It takes $O(n^2 d^4)$ time to compute the d -simplices of $\mathcal{D}(P)$.*

Proof. For every simplex we need to compute $d + 1$ adjacent simplices. This takes $O(nd^3)$ operations. The marks on the $(d - 1)$ -simplices prevent us from searching an already explored d -simplex twice. Therefore, the expansion of each d -simplex happens only once.

The operational costs during the search are bounded from below by the costs for the expansion of a simplex. Since there are $O(n)$ d -simplices in $\mathcal{D}(P)$ and each may be found from $d + 1$ adjacent $(d - 1)$ -simplices, the total costs sum up to $O(n^2 d^4)$ \square

5. CONCLUSIONS

In this article, we presented a new algorithm for the computation of the d -dimensional Delaunay simplices for a finite point set in d -dimensional euclidean space. It is competitive in higher dimensions where the running times of known algorithms tend to be bounded from below by the combinatorial complexity of the Delaunay decomposition. Knowing only the d -simplices is no drawback, as all necessary information may

be deduced by computing intersections. The resulting data structure is an abstract simplicial complex and information may be retrieved in form of a Hasse diagram computed in post processing.

REFERENCES

- [1] B. Gärtner. Fast and Robust Smallest Enclosing Balls. *Proceedings of the 7th Annual European Symposium on Algorithms*. Pages 325-338, 1999.
- [2] R. Sibson. Locally equiangular triangulations. *The Computer Journal* Vol.2 (3) Pages 243-245, 1973.
- [3] L. Guibas, D. Knuth, M. Sharir. Randomized incremental construction of Delaunay and Voronoi diagrams. *Algorithmica*, 7, Pages 381-413, 1992.
- [4] B. Joe. Construction of three-dimensional Delaunay triangulations from local transformations. *Comput. Aided Geom. Design* 8, Pages 123-142, 1991.
- [5] J.D. Boissonnat, O. Devillers, S. Hornus. Incremental Construction of the Delaunay Triangulation and the Delaunay Graph in Medium Dimension. *Proceedings of the twenty-fifth annual symposium on Computational geometry*, Pages 208-216, 2009.
- [6] www.qhull.de
- [7] www.cgal.org

ABOUT THE AUTHOR

Martin Hünninger studied Mathematics and Computer Science at the Friedrich Schiller University in Jena/Germany, where he also received his Dr. rer. nat. for his thesis on topological properties of higher dimensional discrete lattices. Afterwards, he developed algorithms for geometrical data analysis in higher dimensions. He currently works as a deputy professor for Theoretical Computer Science at the Westsaxonian Technical University Zwickau/Germany.

CONGRUENT STEWART GOUGH PLATFORMS WITH NON-TRANSLATIONAL SELF-MOTIONS

Georg Nawratil

Vienna University of Technology, Austria

ABSTRACT: It is well known that each Stewart Gough (SG) manipulator, where the platform is congruent with the base (= congruent SG manipulator), has a 2-dimensional translational self-motion, if all legs have equal (non-zero) length. As congruent SG platforms with planar platform and planar base are only special cases of so-called planar affine/projective SG platforms, which were already studied by the author in foregoing publications, we focus on the non-planar case. In this paper we give a geometric characterization of all non-planar congruent SG platforms, which have further self-motions beside the above mentioned translational one. The main result is obtained by means of bond theory.

Keywords: Stewart Gough platform, Self-motion, Bond theory, Wren platform, Schönflies motion, Cylinder of revolution

1. INTRODUCTION

The geometry of a Stewart Gough (SG) platform is given by the six base anchor points M_i with coordinates $\mathbf{M}_i := (A_i, B_i, C_i)^T$ with respect to the fixed system and by the six platform anchor points m_i with coordinates $\mathbf{m}_i := (a_i, b_i, c_i)^T$ with respect to the moving system (for $i = 1, \dots, 6$). Each pair (M_i, m_i) of corresponding anchor points is connected by a SPS-leg, where only the prismatic joint (P) is active and the spherical joints (S) are passive (cf. Fig. 1).

If the geometry of the manipulator is given as well as the leg lengths, the SG platform is generically rigid. But, under particular conditions, the manipulator can perform a n -dimensional motion ($n > 0$), which is called self-motion.

Note that self-motions are also solutions to the still unsolved problem posed by the French Academy of Science for the "Prix Vaillant" of the year 1904, which is also known as Borel Bricard problem (cf. [1], [2], [9]) and reads as follows: "Determine and study all displacements of a rigid body in which distinct points of the body move on spherical paths."

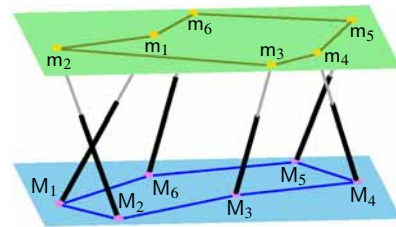


Figure 1: SG manipulator with planar platform and planar base (= planar SG manipulator).

1.1 Bond Theory

In this section we give a short introduction into the theory of bonds for SG manipulators presented in [18], which was motivated by publication [7]. We start with the direct kinematic problem of parallel manipulators of SG type and further with the definition of bonds.

Due to the result of Husty [8], it is advantageous to work with Study parameters ($e_0 : e_1 : e_2 : e_3 : f_0 : f_1 : f_2 : f_3$) for solving the forward kinematics. Note that the first four homogeneous coordinates ($e_0 : e_1 : e_2 : e_3$) are the so-called Euler parameters. Now all real points

of the 7-dimensional Study parameter space P^7 , which are located on the so-called Study quadric $\Psi: \sum_{i=0}^3 e_i f_i = 0$, correspond to an Euclidean displacement, with exception of the 3-dimensional subspace E of Ψ given by $e_0 = e_1 = e_2 = e_3 = 0$, as its points cannot fulfill the condition $N \neq 0$ with $N = e_0^2 + e_1^2 + e_2^2 + e_3^2$. The translation vector $\mathbf{v} := (v_1, v_2, v_3)^T$ and the rotation matrix $\mathbf{R} := (r_{ij})$ of the corresponding Euclidean displacement $\mathbf{R}\mathbf{x} + \mathbf{v}$ are given by:

$$\begin{aligned} v_1 &= 2(e_0 f_1 - e_1 f_0 + e_2 f_3 - e_3 f_2), \\ v_2 &= 2(e_0 f_2 - e_2 f_0 + e_3 f_1 - e_1 f_3), \\ v_3 &= 2(e_0 f_3 - e_3 f_0 + e_1 f_2 - e_2 f_1), \\ r_{11} &= e_0^2 + e_1^2 - e_2^2 - e_3^2, \\ r_{22} &= e_0^2 - e_1^2 + e_2^2 - e_3^2, \\ r_{33} &= e_0^2 - e_1^2 - e_2^2 + e_3^2, \\ r_{12} &= 2(e_1 e_2 - e_0 e_3), \quad r_{21} = 2(e_1 e_2 + e_0 e_3), \\ r_{13} &= 2(e_1 e_3 + e_0 e_2), \quad r_{31} = 2(e_1 e_3 - e_0 e_2), \\ r_{23} &= 2(e_2 e_3 - e_0 e_1), \quad r_{32} = 2(e_2 e_3 + e_0 e_1), \end{aligned}$$

if $N = 1$ is fulfilled. All points of the complex extension of P^7 , which cannot fulfill this normalizing condition, are located on the so-called exceptional cone $N = 0$ with vertex E .

By using the Study parametrization of Euclidean displacements the condition that the point m_i is located on a sphere centered in M_i with radius R_i , is a quadratic homogeneous equation according to Husty [8]. This so-called sphere condition Λ_i has the following form:

$$\begin{aligned} \Lambda_i: \quad & (a_i^2 + b_i^2 + c_i^2 + A_i^2 + B_i^2 + C_i^2 - R_i^2)N + \\ & 2[(a_i A_i + b_i B_i - c_i C_i)e_3^2 - (a_i A_i + b_i B_i + c_i C_i)e_0^2 \\ & - (a_i A_i - b_i B_i - c_i C_i)e_1^2 + (a_i A_i - b_i B_i + c_i C_i)e_2^2 \\ & + 2(c_i B_i - b_i C_i)e_0 e_1 + 2(a_i - A_i)(e_0 f_1 - e_1 f_0) \\ & - 2(c_i A_i - a_i C_i)e_0 e_2 + 2(b_i - B_i)(e_0 f_2 - e_2 f_0) \\ & + 2(b_i A_i - a_i B_i)e_0 e_3 + 2(c_i - C_i)(e_0 f_3 - e_3 f_0) \\ & - 2(b_i A_i + a_i B_i)e_1 e_2 + 2(a_i + A_i)(e_3 f_2 - e_2 f_3) \\ & - 2(c_i A_i + a_i C_i)e_1 e_3 + 2(b_i + B_i)(e_1 f_3 - e_3 f_1) \\ & - 2(c_i B_i + b_i C_i)e_2 e_3 + 2(c_i + C_i)(e_2 f_1 - e_1 f_2) \\ & + 2(f_0^2 + f_1^2 + f_2^2 + f_3^2)] = 0. \end{aligned} \quad (1)$$

Now the solution of the direct kinematics over \mathbb{C} can be written as the algebraic variety V of the ideal \mathcal{I} spanned by $\Psi, \Lambda_1, \dots, \Lambda_6, N = 1$. In general V consists of a discrete set of points with a maximum of 40 elements.¹

We consider the algebraic motion of the mechanism, which are the points on the Study quadric that the constraints define; i.e. the common points of the seven quadrics $\Psi, \Lambda_1, \dots, \Lambda_6$. If the manipulator has a n -dimensional self-motion then the algebraic motion also has to be of this dimension. Now the points of the algebraic motion with $N \neq 0$ equal the kinematic image of V . But we can also consider the points of the algebraic motion, which belong to the exceptional cone $N = 0$. An exact mathematical definition of these so-called bonds can be given as follows (cf. Remark 5 of [18]):

Definition 1 For a SG manipulator the set \mathcal{B} of bonds is defined as:

$$\mathcal{B} := \text{ZarClo}(V^*) \cap \{(e_0 : \dots : f_3) \in P^7 \mid \Psi, \Lambda_1, \dots, \Lambda_6, N = 0\},$$

where V^* denotes the variety V after the removal of all components, which correspond to pure translational motions. Moreover $\text{ZarClo}(V^*)$ is the Zariski closure of V^* , i.e. the zero locus of all algebraic equations that also vanish on V^* .

We have to restrict to non-translational motions for the following reason: A component of V , which corresponds to a pure translational motion, is projected to a single point O (with $N \neq 0$) of the Euler parameter space P^3 by the elimination of f_0, \dots, f_3 . Therefore the intersection of O and $N = 0$ equals \emptyset . Clearly, the kernel of this projection equals the group of translational motions. Moreover it is important to note that the set of bonds only depends on the geometry of the manipulator and not on the leg lengths (cf. Theorem 1 of [18]). For more details please see [18].

¹ Note that for non-planar congruent SG manipulators this number drops to 24 according to [12] and [15].

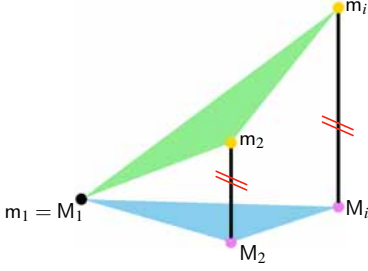


Figure 2: Illustration of Eq. (2) with $m_1 = M_1$.

Due to Theorem 2 of [18] a SG platform possesses a pure translational self-motion if and only if the platform can be rotated about the center $m_1 = M_1$ into a pose, where the vectors $\overrightarrow{M_i m_i}$ for $i = 2, \dots, 6$ fulfill the condition (cf. Fig. 2):

$$rk(\overrightarrow{M_2 m_2}, \dots, \overrightarrow{M_6 m_6}) \leq 1. \quad (2)$$

Moreover all 1-dimensional self-motions are circular translations, which can easily be seen by considering a normal projection of the SG manipulator in direction of the parallel vectors $\overrightarrow{M_i m_i}$ for $i = 2, \dots, 6$. If all these five vectors are zero-vectors, the platform and the base are congruent and therefore we get a so-called congruent SG manipulator. This type of SG manipulator has a well known 2-dimensional translational self-motion \mathcal{T} , if all legs have equal (non-zero) length.

Note that \mathcal{T} is the only 2-dimensional translational self-motion and that higher-dimensional translational self-motions do not exist (cf. [20]).

2. PRELIMINARY CONSIDERATIONS ON CONGRUENT SG PLATFORM

In this article we are interested in designs of congruent SG platforms, which can perform additional self-motions beside \mathcal{T} . First of all we clarify whether congruent SG platforms can have further translational self-motions beside \mathcal{T} . Clearly, due to the result given in the last paragraph of Section 1.1, these translational self-motions can only be 1-dimensional ones.

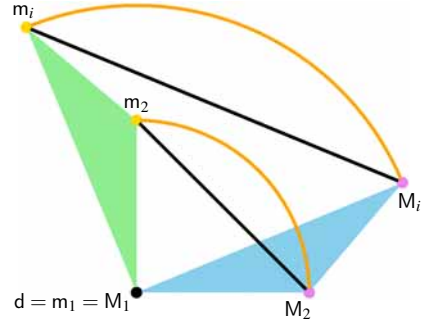


Figure 3: Projection in direction of the axis d : The chords of the rotation around d are parallel if and only if d equals the line of intersection of the planar platform and the planar base.

Lemma 1 *A non-planar congruent SG platform cannot have a 1-dimensional translational self-motion.*

PROOF: If a congruent SG manipulator has a 1-dimensional translational self-motion there has to exist an orientation of the platform with $m_1 = M_1$ and $rk(\overrightarrow{M_2 m_2}, \dots, \overrightarrow{M_6 m_6}) = 1$. We assume that the manipulator is in this configuration.

As the platform and the base are congruent there exists a rotation ρ around the axis d through $m_1 = M_1$ with $\overrightarrow{M_i \rho} = \overrightarrow{m_i}$ for $i = 2, \dots, 6$. Therefore the vectors $\overrightarrow{M_i m_i}$ are chords of ρ . A projection in direction of the axis d shows immediately the validity of this lemma (cf. Fig. 3). \square

For our study we can focus on the non-planar case, as planar congruent SG platforms were already discussed in detail by Karger [14] and by the author [17] as special cases of so-called planar projective SG platforms. Therefore it remains to determine all non-planar congruent SG platforms, which possess non-translational self-motions. This is done in the remainder of this article, which is structured as follows:

In Section 2.1 we discuss non-planar congruent SG platforms, which are architecturally singular. In Section 2.2 we demonstrate on the

basis of a modified Wren platform that non-architecturally singular congruent SG platforms can have non-translational self-motions. After a short review on cylinders of revolution in Section 3, we present a remarkable geometric characterization of all non-planar congruent SG platforms with non-translational self-motions by means of bond theory in Section 4. We close the paper by discussing some non-planar congruent SG platforms with remarkable self-motions in Section 5.

2.1 Architecture singularity

A SG platform is called architecturally singular if it is singular in every possible configuration. All manipulators which have this property are well studied and classified (for a review on this topic see Section 3.1 of [19]). Therefore the following lemma can easily be proven:

Lemma 2 *A non-planar congruent SG platform is architecturally singular if and only if four anchor points are collinear. These manipulators possess self-motions in each pose over \mathbb{C} .*

PROOF: The first sentence follows directly from the list of non-planar architecturally SG platforms given by Karger in Theorem 3 of [13].

If four anchor points are collinear, the corresponding four legs belong to a regulus and one can remove any of the four legs without changing the direct kinematics (\Rightarrow redundant SG platform). This already proves the second statement of Lemma 2. \square

Until now only a few non-architecturally singular SG platforms with self-motions are known. A detailed review of this topic was given in [16], which is as complete as possible to the best knowledge of the author. Based on the body of literature cited within this review, it is not difficult to come up with the following example of a non-planar congruent SG platform with no four anchor points collinear, which possesses a non-translational self-motion.

2.2 Example: Wren platform

We start with considering a special planar congruent SG manipulator, which is also known as Wren platform (cf. [23]). As in this case the anchor points are located on a circle (see Fig. 4, left), the Wren platform is an architecturally singular manipulator (cf. [2], [3], [14]). If all legs have equal length, there also exists a 1-dimensional Schönflies self-motion (see Fig. 5, left), beside \mathcal{T} (see Fig. 6, left). In the left picture of Fig. 4, the *branching singularity* (cf. [6]) of these two self-motions is displayed.

Remark 1 *Due to Wohlhart [23], the Wren platform is called kinematotropic, as it can change the dimension of mobility.* \diamond

Now we consider the Wren platform in its branching configuration (cf. Fig. 4, left). If we translate each leg (including their anchor points) arbitrarily in direction of its carrier line, we end up with a configuration of the modified Wren platform, which is still a congruent SG manipulator, but not longer planar and therefore not architecturally singular (cf. Fig. 4, right).

Clearly the modified Wren platform also has the self-motion \mathcal{T} (cf. Fig. 6, right). Moreover due to Husty and Karger [10] this modification has no influence on the Schönflies self-motion (cf. Fig. 5, right). Therefore this is an example of a non-translational self-motion of a non-planar congruent SG platform, which is not architecturally singular. Note that this existence is not self-evident, as planar congruent SG platforms can only have translational self-motions if they are not architecturally singular (cf. [17]).

Remark 2 *The modified Wren platform also demonstrates that the property of kinematotropy is not restricted to architecturally singular manipulators.* \diamond

Motivated by this example, we are interested in all congruent SG manipulators with non-translational self-motions. Before they are determined in Section 4, we review some known results about cylinders of revolution.

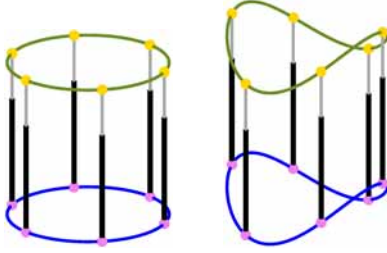


Figure 4: Branching singularity of the Wren platform (left) and the modified one (right).

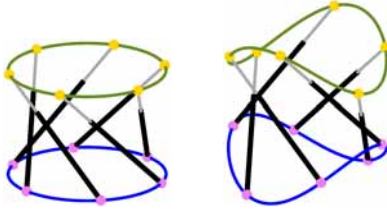


Figure 5: Schönflies self-motion of the Wren platform (left) and the modified one (right).

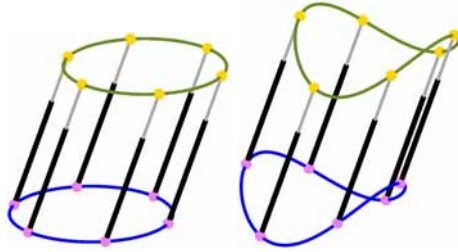


Figure 6: Self-motion \mathcal{T} of the Wren platform (left) and the modified one (right).

3. CYLINDERS OF REVOLUTION

A cylinder of revolution Φ equals the set of all points, which have equal distance to its rotation axis s (finite line). Under the assumption that Φ has at least one real point, we can distinguish the following four cases:

1. s is real and Φ is not reducible: Φ is a cylinder of revolution over \mathbb{R} .
2. s is real and Φ is reducible: Φ equals a pair of isotropic planes² γ_1 and γ_2 , which are conjugate complex. Trivially s carries the only real points of Φ .
3. s is imaginary and Φ is not reducible: Φ is a cylinder of revolution over \mathbb{C} . The real points of Φ are located on the 4th order intersection curve of Φ and its conjugate $\bar{\Phi}$.
4. s is imaginary and Φ is reducible: In this case Φ equals a pair of isotropic planes γ_1 and γ_2 , which are not conjugate complex. Moreover Φ contains two real lines g_i ($i = 1, 2$), which are the intersections of γ_i and its isotropic conjugate $\bar{\gamma}_i$.

Remark 3 *It is a well known fact from projective geometry that the axis s is the line, where the tangent planes γ_1 and γ_2 through s onto Φ are isotropic planes.* \diamond

3.1 Computation of cylinders of revolution

In this section we focus on the determination of all cylinders of revolution through a given set of real points X_1, \dots, X_n . There exist many papers on this well studied problem (see e.g. [5], [22], [24] and the references therein).

In the following we want to use the computational approach of Schaal [22], which was furthered by Zsombor-Murray and El Fashny in [24]. They pointed out that this problem is equivalent with the solution of the following system of equations if X_1 equals the origin U of the reference frame:

$$s^2 = 1, \quad (3)$$

$$\Upsilon: \quad s \cdot t = 0, \quad (4)$$

$$\Omega_i: \quad (\mathbf{x}_i \times \mathbf{s})^2 - 2s^2(\mathbf{x}_i \cdot \mathbf{t}) = 0, \quad (5)$$

for $i = 2, \dots, n$, where \mathbf{x}_i is the coordinate vector of the point X_i , $\mathbf{s} := (s_1, s_2, s_3)^T$ the direction

²A plane is called isotropic if its ideal line is tangent to the absolute quadric.

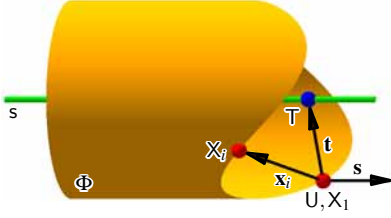


Figure 7: Notation used for computation.

vector of the rotation axis s , and $\mathbf{t} := (t_1, t_2, t_3)^T$ the coordinate vector of the footpoint T on s with respect to $U = X_1$ (cf. Fig. 7).

The rough procedure for solving this system of equations is as follows: In the first step, one solves the equations $\Upsilon, \Omega_2, \dots, \Omega_n$, which already gives the solutions up to a common factor; i.e. we get $s_1 : s_2 : s_3 : t_1 : t_2 : t_3$. In the second step, we normalize these 6-tuples with respect to the normalizing condition given in Eq. (3). This normalization is always possible as the axis cannot be isotropic³, because it is the intersection of two isotropic planes (cf. Remark 3).

Remark 4 For $n = 5$ there exist in general six cylinders of revolution over \mathbb{C} (e.g. [24]). There even exist examples, where all six cylinders are real (e.g. [5]). For $n > 5$ no solution exists, if X_1, \dots, X_n are in general configuration. \diamond

4. MAIN THEOREM

Based on these considerations regarding cylinders of revolution, we can formulate the following main theorem:

Theorem 1 *A non-planar congruent SG manipulator can have a real non-translational self-motion only if the six base (resp. platform) anchor points have equal distance to a finite line s , i.e. they are located on a cylinder of revolution of type 1, 3 or 4 listed in Section 3. Moreover this condition is also sufficient for the existence of self-motions over \mathbb{C} .*

³The line is called isotropic if its ideal point is located on the absolute quadric.

PROOF: The proof of this theorem is based on the following fact: If a non-translational self-motion exists, then the bond-set has to be non-empty. Therefore we have to determine the conditions for which the set of bonds consists of at least one element. The computation of these conditions is outlined next.

Without loss of generality (w.l.o.g.) we can choose Cartesian coordinate systems in the platform and base with $a_i = A_i, b_i = B_i, c_i = C_i$ for $i = 1, \dots, 6$ and $a_1 = b_1 = b_2 = c_1 = c_2 = c_3 = 0$. In order to eliminate the factor of similarity we can set $a_2 = 1$. Clearly we assume for the remainder of this article that all anchor points are distinct, as otherwise two legs coincide due to the congruence of the platform and the base. Moreover we can assume (after a possible necessary reindexing of anchor points) that the first four points are not coplanar; i.e. $b_3c_4 \neq 0$. We distinguish three cases.

4.1 No three anchor points are collinear

According to [18] the set of bonds can be computed as follows: We compute $\Delta_{j,i} := \Lambda_j - \Lambda_i$, which is only linear in the Study parameters f_0, \dots, f_3 . We distinguish three cases.

General case $e_3 \neq 0$. Assuming a real motion we can solve the linear system of equations $\Psi, \Delta_{2,1}, \Delta_{3,1}$ for f_1, f_2, f_3 without loss of generality. We plug the obtained expressions for f_1, f_2, f_3 into $\Delta_{4,1}, \Delta_{5,1}, \Delta_{6,1}$ and consider their numerators, which are homogeneous polynomials P_4, P_5, P_6 of degree three in the Euler parameters. Note that these polynomials do not depend on f_0 . Therefore f_0 can be calculated from Λ_1 , but this is not of interest for the further computation.

We eliminate e_0 from P_i and $N = 0$ by computing the resultant Q_i of these two expressions for $i = 4, 5, 6$. Q_i factors into $16F_i^2$ with

$$F_i = \sum_{j+k+l=3} g_{jkl} e_1^j e_2^k e_3^l \quad \text{and} \quad j, k, l \in \{0, \dots, 3\}.$$

Moreover the coefficients g_{jkl} are given by:

$$\begin{aligned} g_{210} &= c_i b_3^2, & g_{201} &= b_3(b_i^2 - b_3 b_i + c_i^2), \\ g_{300} &= 0, & g_{012} &= c_i(b_3^2 + a_3^2 - a_3 - 2b_3 b_i), \\ g_{120} &= b_3 c_i(1 - 2a_3), & g_{102} &= b_3 c_i(1 - 2a_i), \\ g_{111} &= 2b_3 b_i(a_3 - a_i), & g_{030} &= a_3 c_i(a_3 - 1), \\ g_{021} &= a_i^2 b_3 - a_i b_3 + a_3 b_i - a_3^2 b_i + c_i^2 b_3, \\ g_{003} &= a_3 b_i - a_3^2 b_i + a_i^2 b_3 + b_i^2 b_3 - a_i b_3 - b_i b_3^2. \end{aligned}$$

We proceed with the computation of the resultant U_k of F_i and F_j with respect to e_1 for pairwise distinct $i, j, k \in \{4, 5, 6\}$. From each U_k we can factor out $b_3^2 e_3^2$ and remain with an expression V_k with 3827 terms. Moreover V_k is a homogeneous polynomial of degree 6 in the remaining Euler parameters e_2, e_3 .

Now the necessary condition for the existence of a bond is that V_4, V_5, V_6 have a common solution; i.e. the resultant W_k of V_i and V_j with respect to e_2 with pairwise distinct $i, j, k \in \{4, 5, 6\}$, has to be fulfilled identically for all e_3 . Due to the large number of terms and the high degree, W_k cannot be computed in general, and therefore it seems that we cannot prove the theorem.

But due to the example of the modified Wren platform, we conjecture that bonds can only exist if the six anchor points are located on a cylinder of revolution. Therefore we consider the system of equations $\Upsilon, \Omega_2, \dots, \Omega_6$ given in Eqs. (4) and (5) with respect to the six anchor points.

Under the assumption $s_3 \neq 0$, we can solve $\Upsilon, \Omega_2, \Omega_3$, which are linear in t_1, t_2, t_3 , for these unknowns. We plug the obtained expressions into $\Omega_4, \Omega_5, \Omega_6$ and consider their numerators, which are homogeneous polynomials F_4^*, F_5^*, F_6^* . For the substitution $s_i \leftrightarrow e_i$ for $i = 1, 2, 3$ the polynomial F_j^* equals F_j for $j = 4, 5, 6$.

Therefore the existence of a cylinder of revolution through the six anchor points implies the existence of a bond and vice versa. This already closes the general case.

Special case $e_3 = 0, e_2 \neq 0$. The procedure for this case only differs slightly from the general one. Assuming a real motion we can solve

$\Psi, \Delta_{2,1}, \Delta_{4,1}$ for f_1, f_2, f_3 without loss of generality. Then we plug the obtained expressions into $\Delta_{3,1}, \Delta_{5,1}, \Delta_{6,1}$ and consider their numerators, which are homogeneous polynomials P_3, P_5, P_6 of degree two in the Euler parameters. We eliminate e_0 from P_i and $N = 0$ by computing the resultant Q_i of these two expressions for $i = 3, 5, 6$. Q_i factors into $16e_2^2 F_i^2$ with $F_i = g_{20}e_1^2 + g_{11}e_1e_2 + g_{02}e_2^2$ and

$$\begin{aligned} g_{20} &= c_i c_4^2 + c_i b_4^2 - b_i^2 c_4 - c_i^2 c_4, \\ g_{11} &= 2a_i b_i c_4 - b_i c_4 + c_i b_4 - 2c_i a_4 b_4, \\ g_{02} &= c_i a_4^2 - c_i a_4 - c_i^2 c_4 + a_i c_4 + c_i c_4^2 - a_i^2 c_4. \end{aligned}$$

Now we proceed with the computation of the cylinders of revolution with $s_3 = 0$ and $s_2 \neq 0$. W.l.o.g. we can solve $\Upsilon, \Omega_2, \Omega_4$ for t_1, t_2, t_3 . We plug the obtained expressions into $\Omega_3, \Omega_5, \Omega_6$ and consider their numerators, which are homogeneous polynomials F_3^*, F_5^*, F_6^* . Again the substitution $s_i \leftrightarrow e_i$ for $i = 1, 2$ shows that F_j^* equals F_j for $j = 3, 5, 6$. Therefore we can draw the same conclusion as in the general case.

Very special case $e_2 = e_3 = 0$. If $e_1 = 0$ holds, the platform has the same orientation during the whole self-motion. As a consequence we can only end up with a translational self-motion, which has to be 2-dimensional due to Lemma 1.

Assuming a real motion with $e_1 \neq 0$ we can solve $\Psi, \Delta_{3,1}, \Delta_{4,1}$ for f_1, f_2, f_3 without loss of generality. Then we plug the obtained expressions into $\Delta_{2,1}, \Delta_{5,1}, \Delta_{6,1}$ and consider their numerators, which are homogeneous polynomials P_2, P_5, P_6 of degree two in the Euler parameters. Note that P_2 factors into $N(R_1 - R_2)$. Therefore we only have to compute the resultant Q_i of $N = 0$ and P_i with respect to e_0 for $i = 5, 6$. Q_i factors into $16b_3^2 F_i^2$ with $F_i = g_2 e_1^2$ and

$$g_2 = b_3 c_i b_4 + b_i^2 c_4 + c_i^2 c_4 - c_i c_4^2 - c_3 c_i c_4 - c_i b_4^2.$$

Now we proceed with the computation of the cylinders of revolution with $s_2 = s_3 = 0$ and $s_1 \neq 0$. W.l.o.g. we can solve $\Upsilon, \Omega_3, \Omega_4$ for

t_1, t_2, t_3 . If we plug the obtained expression into Ω_2 , we see that it is fulfilled identically. Therefore we consider the numerators of Ω_5, Ω_6 , which are homogeneous polynomials F_5^*, F_6^* . Again the substitution $s_1 \leftrightarrow e_1$ shows that F_j^* equals F_j for $j = 5, 6$. Therefore we can draw the same conclusion as in the general case.

Remark 5 One also has to check in the general, special and very special case that Q_i can always be computed by means of resultant. This is true if the coefficient H_i of e_0^2 in P_i does not vanish.⁴ As the bonds are independent of the leg lengths, H_i has to vanish independently from R_1, \dots, R_6 . It can easily be seen that this cannot be the case without contradicting our assumptions. \diamond

This closes the study of cases, where no three anchor points are collinear. Clearly in this case we can only obtain cylinders of revolution of type 1 and 3 of the list given in Section 3.

4.2 Three anchor points are collinear but no four anchor points are collinear

We assume that the first, second and fifth anchor points are collinear, which implies $b_5 = c_5 = 0$. We distinguish again the following three cases:

General case $e_3 \neq 0$. Everything can be done similarly to the general case of Section 4.1.

Special case $e_3 = 0, e_2 \neq 0$. Everything can be done similarly to the special case of Section 4.1.

Very special case $e_2 = e_3 = 0$. Everything can be done similarly to the very special case of Section 4.1. It should only be noted that in this case also P_5 factors into $N(R_1 - R_5)$. As a consequence not only Ω_2 is fulfilled identically in the corresponding computation of the cylinders of revolution, but also Ω_5 .

Due to the collinearity of three anchor points and the exclusion of architecturally singular designs, we can only obtain cylinders of revolution

⁴There are no terms with e_0^3 and e_0 in P_i .

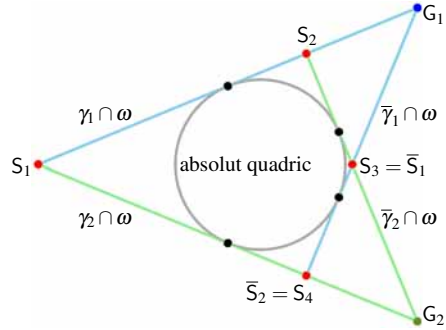


Figure 8: Sketch of the situation in the ideal plane ω : G_i denotes the ideal point of the line g_i for $i = 1, 2$. Moreover S_1, \dots, S_4 denote the ideal points of the axes s_1, \dots, s_4 of the four different cylinders of revolution of type 4. Note that the first and third solution as well as the second and fourth one are conjugate complex.

of type 4 of the list given in Section 3 if the carrier line of the three collinear points is no generator of Φ (general case and special one). Then the six points have to be located on two skew lines g_1 and g_2 , where each line carries three pairwise distinct anchor points. For this non-planar congruent SG manipulator there always exists four cylinders of revolution of type 4 (cf. Section 5.1), which can easily be seen on basis of Fig. 8.

In the very special case, the carrier line of the three collinear points is a generator of Φ , thus we can only obtain a cylinder of revolution of type 1 of the list given in Section 3.

4.3 Four anchor points are collinear

As the anchor points are located on two skew lines, there also exist the above mentioned four cylinders of revolution of type 4 (cf. Fig. 8). Moreover there is one cylinder of revolution Φ of type 1, where the carrier line of the four collinear anchor points is a generator of Φ . The cylinder Φ is of multiplicity two, thus in sum this redundant manipulator (cf. Proof of Lemma 2) has the expected six solutions (cf. Remark 4).

The existence of bonds follows trivially from the second part of Lemma 2.

Remark 6 *With exception of this architecturally singular case, there are no examples known to the author, where the anchor points are located on more than four cylinder of revolution (cf. Sections 5.1 and 5.2).* \diamond

4.4 Sufficiency

In this section we want to point out that the condition that the six points are located on a cylinder of revolution of type 1, 3 or 4 is already sufficient for the existence of a self-motion over \mathbb{C} . Geometrically this is clear, as for each cylinder of revolution there exists the corresponding Schönflies self-motion, which was recognized for the modified Wren platform (cf. Section 2.2). This circumstance can also easily be verified algebraically as follows:

As for this Schönflies self-motion all legs have to have equal length, we set $R_1 = \dots = R_6$. Then we do not have to eliminate e_0 by applying the resultant with $N = 0$, but P_i already equals F_i up to a non-zero factor. Therefore e_0 remains free and parametrizes the self-motion. This finishes the proof of Theorem 1. \square

5. SELF-MOTIONS

From the algebraic proof of the sufficiency given in Section 4.4, it also follows that the Schönflies self-motion is real if and only if the points are located on a real cylinder of revolution, as the direction $(s_1 : s_2 : s_3)$ of the axis s equals the direction $(e_1 : e_2 : e_3)$ of the rotation axis of the Schönflies self-motion.

But we cannot conclude from Theorem 1 that the Schönflies self-motions are the only non-translational self-motions, which can be performed by the manipulators characterized in Theorem 1. A trivial counter example is the architecture singularity, as the self-motions are the motions of the 5-legged manipulator, which results from the removal of one of the four legs, whose anchor points are collinear (cf. Lemma 2).

Further counter examples read as follows:

5.1 Butterfly self-motions

If M_1, M_2, M_3 are collinear and M_4, M_5, M_6 are collinear, then there also exist the following self-motion: If the platform is placed in a way that m_4, m_5, m_6 (resp. m_1, m_2, m_3) are located on the carrier line g of M_1, M_2, M_3 (resp. M_4, M_5, M_6), then we get a pure rotational self-motion about g , which is called butterfly motion.

It is already known (cf. Examples 1 and 2 of [18]) that a Schönflies self-motion possesses only one bond (up to conjugation of coordinates), which is singular⁵, and that a butterfly self-motion is given by two 1-parametric bonds (up to conjugation of coordinates).

As there are two butterfly self-motions we get four 1-parametric bonds (up to conjugation of coordinates). They possess four singular bonds (up to conjugation of coordinates), which correspond with the four cylinders of revolution of type 4 (cf. Fig. 8) implying four complex Schönflies self-motions.

5.2 New self-motions

We study non-planar congruent SG platforms which are plane-symmetric, i.e. the fourth, fifth and sixth anchor point are obtained by reflecting the first, second and third one on a plane ε . Therefore there always exists a cylinder of revolution Φ of type 1 with generators orthogonal to ε .

W.l.o.g. we can assume that ε is the xy -plane and that the rotation axis of Φ is the z -axis. Moreover we can eliminate the factor of similarity by setting the radius of Φ equal to 1. Finally we can rotate the coordinate system about the z -axis in a way that the i -th and j -th anchor point have the same y -coordinate, which results in the following coordinatization:

$$\begin{aligned} a_i &= a_{i+3} = \sin(\mu), & b_i &= b_{i+3} = \cos(\mu), \\ a_j &= a_{j+3} = \sin(-\mu), & b_j &= b_{j+3} = \cos(\mu), \\ a_k &= a_{k+3} = \sin(\lambda), & b_k &= b_{k+3} = \cos(\lambda), \end{aligned}$$

$$c_i = -c_{i+3} \neq 0, c_j = -c_{j+3} \neq 0, c_k = -c_{k+3} \neq 0$$

⁵A bond is called singular if it is located in the vertex space E of $N = 0$.

with pairwise distinct $i, j, k \in \{1, 2, 3\}$ and the angles $\mu \in (0, \pi)$ and $\lambda \in [0, 2\pi)$.

Due to the plane-symmetry these six anchor points are located on further three cylinders of revolution⁶ beside Φ , whose three axes of rotation are located within ε . Therefore the bond-set of this plane-symmetric congruent SG platform consists (up to conjugation of coordinates) of the four singular bonds of the Schönflies self-motions implied by the four cylinders of revolution. But beside these self-motions there exist the following ones characterized by $e_3 = 0$, which are new to the best knowledge of the author:

As we can assume $e_2 \neq 0$ w.l.o.g.⁷, the unknowns f_1, f_2, f_3 can be computed from $\Psi, \Delta_{j,i}, \Delta_{i+3,i}$. If we plug the obtained expressions into $\Delta_{j+3,i}$, it can easily be seen that it vanishes for

$$R_{j+3}^2 = \frac{c_j}{c_i}(R_{i+3}^2 - R_i^2) + R_j^2.$$

Moreover, if additionally

$$R_{k+3}^2 = \frac{c_k}{c_i}(R_{i+3}^2 - R_i^2) + R_k^2$$

is fulfilled, $\Delta_{k,i} = \Delta_{k+3,i}$ holds. The numerator of this condition is a homogeneous polynomial P_k of degree 3 in the Euler parameters e_0, e_1, e_2 . Note that P_k does not depend on f_0 , which is determined by the only remaining equation $\Lambda_i = 0$. Hence, for given five design parameters $c_i, c_j, c_k, \mu, \lambda$, the cubic P_k implies a 4-parametric set \mathcal{S} of self-motions, as it depends on the four leg lengths R_i, R_j, R_k, R_{i+3} .

Subset \mathcal{X} of \mathcal{S} . For the following special choice of R_j and R_k :

$$R_j^2 = \frac{c_j}{2c_i}(R_i^2 - R_{i+3}^2) + \frac{1}{2}(R_i^2 + R_{i+3}^2),$$

$$R_k^2 = \frac{c_k}{2c_i}(R_i^2 - R_{i+3}^2) + \frac{1}{2}(R_i^2 + R_{i+3}^2),$$

⁶Either all three are real (type 1) or one is real and the remaining two are conjugate complex (type 3).

⁷If $e_2 = 0$ holds, the self-motion is a Schönflies motion, where the rotation axis is parallel to the x -axis. But in this case i, j, k can be permuted (\Rightarrow change of coordinate system) in a way that this parallelism vanishes ($\Rightarrow e_2 \neq 0$).

P_k does not depend on the Euler parameter e_0 and the remaining leg lengths R_i and R_{i+3} . For reasons given later on, the three solutions of the cubic equation $P_k = 0$ with respect to e_1 and e_2 have to equal the axial directions of the three cylinder axes located in ε . Therefore each of these three directions determines a 2-parametric set \mathcal{X} of Schönflies self-motions, as the leg lengths R_i and R_{i+3} can still be chosen arbitrarily. Note that the choice of R_i and R_{i+3} only influences the translational component of the Schönflies self-motion.

For the extra condition $R_i = R_{i+3}$ we obtain the 1-parametric set \mathcal{X}_- of Schönflies self-motions with equal leg lengths $R_1 = \dots = R_6$, which were recognized for the modified Wren platform (cf. Section 2.2). This proves the validity of the above given solutions of $P_k = 0$, as this cubic equation (and therefore its solutions) does not depend on R_i and R_{i+3} .

Remark 7 Due to the limitation of length, a discussion of exemplary self-motions generated by a plane-symmetric congruent SG platform can be downloaded from the author's homepage. Furthermore animations of these self-motions are provided as supplementary data. \diamond

5.3 Addendum

After finishing this paper the author became aware of Example 3.3.8 given by Husty et al. [11], where a full discussion of the self-motions of the following special congruent SG platform is given:

$$a_1 = -a_2 \neq 0, \quad b_3 = -b_4 \neq 0, \quad c_5 = -c_6 \neq 0,$$

and all remaining coordinates are zero.

Beside the 2-dimensional translational self-motion \mathcal{T} , there exist Schönflies self-motions⁸ and more complicated self-motions of 4th order, which possess exact six points with spherical trajectories. The latter self-motions were firstly reported by Dietmaier in [4] and have the property $R_1 = R_2, R_3 = R_4$ and $R_5 = R_6$.

⁸Due to the symmetry of the manipulator there always exist cylinders of revolution through the six anchor points, which imply the Schönflies self-motions.

6. CONCLUSIONS AND OUTLOOK

The first research on cylinders of revolution through a given non-planar set of points by Schaal [22] was motivated kinematically. After a pure geometric and algebraic study done by several researchers (see e.g. [5], [24] and the references therein), it was shown within the paper at hand that this problem is strongly connected with the following one, which is again kinematic: The determination of all non-planar congruent SG platforms with a real non-translational self-motion. It turns out that the six base (resp. platform) anchor points have to be located on a cylinder of revolution (cf. Theorem 1). Note that this geometric characterization is also sufficient for the existence of self-motions over \mathbb{C} . Based on the modified Wren platform given in Section 2.2, this main theorem was proven by means of bond theory.

Although this result is known, a complete list of all possible non-translational self-motions of congruent SG platforms is still missing. All known examples are given in Section 5, where also a family of new self-motions is presented (cf. Section 5.2). Moreover a restriction of the sufficiency condition with respect to \mathbb{R} remains also open.

In the future we are interested in the generalization of the presented result with respect to the linear coupling κ of the non-planar platform and the base. Based on the paper at hand, this was already done in [21] for the case that κ is a similarity. The problem for the case, where κ is an affinity or even a projectivity, remains open.

ACKNOWLEDGMENTS

This research is funded by Grant No. I 408-N13 and Grant No. P 24927-N25 of the Austrian Science Fund FWF.

REFERENCES

- [1] E. Borel. Mémoire sur les déplacements à trajectoires sphériques. *Mémoire présentées par divers savants à l'Académie des Sciences de l'Institut National de France*, 33(1):1–128, 1908.
- [2] R. Bricard. Mémoire sur les déplacements à trajectoires sphériques. *Journal de École Polytechnique*, 11:1–96, 1906.
- [3] M. Chasles. Sur les six droites qui peuvent être les directions de six forces en équilibre. *Comptes Rendus des Séances de l'Académie des Sciences*, 52:1094–1104, 1861.
- [4] P. Dietmaier. Forward kinematics and mobility criteria of one type of symmetric Stewart-Gough platforms. *Recent Advances in Robot Kinematics (J. Lenarcic, V. Parenti-Castelli eds.)*, 379–388, Kluwer, 1996.
- [5] A. Gferrer and P. Zsombor-Murray. Quadrics of Revolution on Given Points. *Journal for Geometry and Graphics*, 13(2):131–144, 2009.
- [6] G. Gogu. Branching singularities in kinematotropic parallel mechanisms. *Computational Kinematics (A. Kecskemethy, A. Müller eds.)*, 341–348, Springer, 2009.
- [7] G. Hegedüs, J. Schicho, and H-P. Schröcker. Bond Theory and Closed 5R Linkages. *Latest Advances in Robot Kinematics (J. Lenarcic, M. Husty eds.)*, 221–228, Springer, 2012.
- [8] M. Husty. An algorithm for solving the direct kinematics of general Stewart-Gough platforms. *Mechanism and Machine Theory*, 31(4):365–380, 1996.
- [9] M. Husty. E. Borel's and R. Bricard's Papers on Displacements with Spherical Paths and their Relevance to Self-Motions of Parallel Manipulators. *International Symposium on History of Machines and Mechanisms (M. Ceccarelli ed.)*, 163–172, Kluwer, 2000.
- [10] M. Husty and A. Karger. Self motions of Stewart-Gough platforms: an overview.

- Proc. of the workshop on fundamental issues and future research directions for parallel mechanisms and manipulators (C.M. Gosselin, I. Ebert-Uphoff eds.)*, 131–141, 2002.
- [11] M. Husty, A. Karger, H. Sachs, and W. Steinhilper. *Kinematik und Robotik*. Springer, 1997.
 - [12] C. Innocenti. Forward kinematics of a 6-6 fully-parallel manipulator with congruent base and platform. *Advances in Robot Kinematics: Analysis and Control (J. Lenarcic, M.L. Husty eds.)*, 137–146, Kluwer, 1998.
 - [13] A. Karger. Architecturally singular non-planar parallel manipulators. *Mechanism and Machine Theory*, 43(3):335–346, 2008.
 - [14] A. Karger. Singularities and self-motions of a special type of platforms. *Advances in Robot Kinematics: Theory and Applications (J. Lenarcic, F. Thomas eds.)*, 155–164, Springer, 2002.
 - [15] C. Mavroidis. Completely specified displacements of a rigid body and their application in the direct kinematics of in-parallel mechanisms. *Proc. of the ASME Design Engineering Technical Conference*, 1998.
 - [16] G. Nawratil. Review and recent results on Stewart Gough platforms with self-motions. *Applied Mechanics and Materials*, 162:151–160, 2012.
 - [17] G. Nawratil. Self-motions of planar projective Stewart Gough platforms. *Latest Advances in Robot Kinematics (J. Lenarcic, M. Husty eds.)*, 27–34, Springer, 2012.
 - [18] G. Nawratil. Introducing the theory of bonds for Stewart Gough platforms with self-motions. *ASME Journal of Mechanisms and Robotics*, 6(1):011004, 2014.
 - [19] G. Nawratil. Correcting Duporcq’s theorem. *Mechanism and Machine Theory*, 73:282–295, 2014.
 - [20] G. Nawratil. On Stewart Gough manipulators with multidimensional self-motions. *Computer Aided Geometric Design*, accepted
 - [21] G. Nawratil. On equiform Stewart Gough manipulators with self-motions. *Journal for Geometry and Graphics*, 17(2):163–175, 2013.
 - [22] H. Schaal. Ein geometrisches Problem der metrischen Getriebesynthese. *Sitzungsber. österr. Akad. Wiss., Math.-Naturw. Kl., Abt. II*, 194(1-3):39–53, 1985.
 - [23] K. Wohlhart. Kinematotropic linkages. *Recent Advances in Robot Kinematics (J. Lenarcic, V. Parenti-Castelli eds.)*, 359–368, Kluwer, 1996.
 - [24] P. Zsombor-Murray and S. El Fashny. A Cylinder of Revolution on Five Points. *Journal for Geometry and Graphics*, 10(2):207–213, 2006.

ABOUT THE AUTHOR

Georg Nawratil is a senior researcher in the research group “Differential Geometry and Geometric Structures” at the Institute of Discrete Mathematics and Geometry, Vienna University of Technology, Austria. His research interests include the geometry of mechanisms, kinematics, robotics and line geometry. He can be reached by email: nawratil@geometrie.tuwien.ac.at, by phone: +43-1-58801-104362, by fax: +43-1-58801-9104362, or through the postal address: Institute of Discrete Mathematics and Geometry, Vienna University of Technology, Wiedner Hauptstrasse 8-10/104, A-1040 Vienna, Austria, Europe. More informations about the author can be found on his homepage: <http://www.geometrie.tuwien.ac.at/nawratil/>.

CONTACT OF THE RULED NONDEVELOPABLE SURFACES

Konstantin Leonidovich PANCHUK and Anton Sergeevich NITEYSKIY
Omsk State Technical University, Russia

ABSTRACT: In the theory of ruled surfaces are known researches on the contact of ruled surfaces along their common generator. In known monograph "Computational Line Geometry" of H. Pottmann and J. Wallner for these purposes Klein image is used. In this paper we propose a study of contact of nondevelopable ruled surfaces via the dual vector calculus. The advantages of this method have been demonstrated by W. Blaschke and D. N. Zeiliger in differential geometry studies of ruled surfaces in space R^3 over the algebra of dual numbers.

Keywords: Ruled nondevelopable surfaces, Dual vector calculus, Order of contact.

1. ELEMENTS OF THEORY

According to these mathematical tools, the equation of a ruled surface can be represented as a unit dual vector-function with a real parameter t :

$$A_1(t) = a_{01}(t) + \omega a_{11}(t), |A_1(t)| = 1, \omega^2 = 0,$$

where $a_{01}(t)$ - a unit generator vector of ruled surface, $a_{11}(t)$ - a momentum vector of generator with respect to the origin. $A_1(t)$ is $n+1$ times continuously differentiable. It's also known that trihedron of the ruled surface consisting of dual unit vectors:

$$A_1; A_2 = \frac{A_1'}{|A_1'|}; A_3 = A_1 \times A_2,$$

where $A_1' = \frac{dA_1}{dt}$; $H_A = |A_1'|$; A_2 - a central normal's vector; A_3 - a center tangent's vector. And is also used an inner geometric parameter of a ruled surface - an element of dual arc:

$ds_A = ds_{A0} + \omega \cdot ds_{A1} = H_A dt, \omega^2 = 0$, which means a dual angle between its two infinitesimally close generators. From an element of dual arc the distribution parameter of the ruled

surface is follow: $p = \frac{ds_{A1}}{ds_{A0}}$.

Trihedron derivation equations of W. Blaschke [2] are of the form: $A_1' = H_A \cdot A_2$;

$$A_2' = -H_A \cdot A_1 + Q_A \cdot A_3; \quad A_3' = -Q_A \cdot A_2,$$

where $H_A = h_{A0} + \omega h_{A1} = |A_1'|$,

$$Q_A = q_{A0} + \omega q_{A1} = \frac{(A_1 A_1')^*}{H_A^2}, \quad \omega^2 = 0.$$

We consider a contact of two nondevelopable ruled surfaces parametrized by $A_1(t)$ and $B_1(t)$ with a common ruling $a_{01}(t) = b_{01}(t)$ at parameter value $t=t_0$. Initially let's introduce the dual vector of divergence $G(t)$ in their common ruling. Because of each unit dual vector function $A_1(t)$ and $B_1(t)$ can be represented as a Taylor series it can be written:

$$G(t) = A_1(t_0) - B_1(t_0) + [A_1'(t_0) - B_1'(t_0)] \cdot \Delta t + \\ + [A_1''(t_0) - B_1''(t_0)] \cdot \frac{\Delta t^2}{2!} + \dots \quad (1)$$

We will define the concept of an order contact of two ruled surfaces in their common generator line. Each of considered dual vector function has a coordinate representation

$$\begin{aligned} A_1(t) &= i \cdot x_A(t) + j \cdot y_A(t) + k \cdot z_A(t), \\ B_1(t) &= i \cdot x_B(t) + j \cdot y_B(t) + k \cdot z_B(t), \end{aligned} \quad (2)$$

where $x_A, y_A, z_A; x_B, y_B, z_B$ – a dual scalar functions. It allows to write an expression for dual vector $G(t)$:

$$G(t) = A_1(t) - B_1(t) = i \cdot x_G + j \cdot y_G + k \cdot z_G = i(x_A - x_B) + j(y_A - y_B) + k(z_A - z_B). \quad (3)$$

Dual vector $G(t)$ characterizes the closeness of ruled surfaces $A_1(t)$ and $B_1(t)$ in infinitesimal neighborhood for their common generator at parameter $t = t_0$. It is defined by two generators $a'_{01}(t)$ and $b'_{01}(t)$ each of which are shifted in their ruled surface on the same dual arc $ds_A = ds_B$ relatively to their common generator $a_{01}(t_0) = b_{01}(t_0)$ (see Fig. 1).

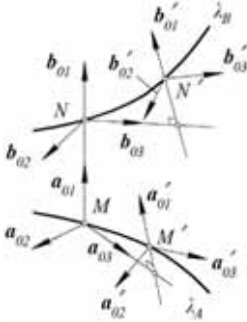


Figure 1: The $n=0$ order contact of the ruled surfaces

Let's introduce the notion of the order contact of two ruled surfaces in their common generator. We will estimate an order of infinitesimality of the module $|G(t)|$ relatively to a dual infinitesimal Δs . If the n – positive integer and following condition is holds

$$\lim_{\Delta s \rightarrow 0} \frac{|G|}{\Delta s^n} = 0, \quad (4)$$

thus we will assume that ruled surfaces in their common generator $a_{01}(t_0) = b_{01}(t_0)$ are not lower than n -th order contact. But if n – maximal then ruled surfaces are in exactly n -th order contact.

Because of $ds = ds_0 + \omega ds_1 = h_0 dt + \omega h_1 dt$, $\omega^2 = 0$, where is $ds_0 = h_0 dt$, $ds_1 = h_1 dt$, so the condition

$\Delta s \rightarrow 0$ can be replaced by following two conditions: $\Delta s_0 \rightarrow 0$ and $\Delta s_1 \rightarrow 0$, that with $h_0 \neq 0$ and $h_1 \neq 0$ leads to $\Delta t \rightarrow 0$. Indeed ds_0 and dt , ds_1 and dt – pairs of equivalent infinitesimals. As a result we obtain the following condition for estimating the order contact of ruled surfaces in their common generator:

$$\lim_{\Delta t \rightarrow 0} \frac{|G(t)|}{\Delta t^n} = 0. \quad (5)$$

Theorem. To fulfill the condition of (5) it is necessary and sufficiently that following equalities are carried out

$$\begin{aligned} x_G(t_0) &= 0, x'_G(t_0) = 0, \\ x''_G(t_0) &= 0, \dots, x_G^{(n)}(t_0) = 0; \\ y_G(t_0) &= 0, y'_G(t_0) = 0, \\ y''_G(t_0) &= 0, \dots, y_G^{(n)}(t_0) = 0; \\ z_G(t_0) &= 0, z'_G(t_0) = 0, \\ z''_G(t_0) &= 0, \dots, z_G^{(n)}(t_0) = 0; \end{aligned} \quad (6)$$

Sufficiency. Assume that the equalities (6) are holds. Then from (1) follows:

$$\begin{aligned} x_G(t) &= x_A(t) - x_B(t) = x^{(n+1)}(\delta_1) \cdot \frac{\Delta t^{n+1}}{(n+1)!}; \\ y_G(t) &= y_A(t) - y_B(t) = y^{(n+1)}(\delta_2) \cdot \frac{\Delta t^{n+1}}{(n+1)!}; \\ z_G(t) &= z_A(t) - z_B(t) = z^{(n+1)}(\delta_3) \cdot \frac{\Delta t^{n+1}}{(n+1)!}, \end{aligned}$$

where $\delta_1, \delta_2, \delta_3$ – values of parameter t which satisfying the conditions: $t_0 < \delta_1 < t$, $t_0 < \delta_2 < t$, $t_0 < \delta_3 < t$. Thus the condition of (5) is holds.

Necessity. Assume that at least one of the equalities (6) is not holding: for example $x_G^{(k)}(t_0) = x_A^{(k)}(t_0) - x_B^{(k)}(t_0) \neq 0$; $k < n$. Thus

$$\begin{aligned} x_G(t) &= x_A(t) - x_B(t) = \\ &= [x_A^{(k)}(t_0) - x_B^{(k)}(t_0)] \cdot \frac{\Delta t^k}{k!} + \dots, \end{aligned}$$

hence

$$\lim_{\Delta t \rightarrow 0} \frac{|G(t)|}{\Delta t^n} \neq 0, k < n.$$

So the condition of (5) is not holds.

Let's consider the primary orders contact of ruled surfaces.

1.1 The n=0 order contact of the ruled surfaces

$$A_1(t_0) = B_1(t_0); A_1'(t_0) \neq B_1'(t_0). \quad (7)$$

In this case ruled surfaces are intersecting along their common generator (see Fig. 1).

1.2 The n=1 order contact of the ruled surfaces

$$A_1(t_0) = B_1(t_0); A_1'(t_0) = B_1'(t_0); \\ A_1''(t_0) \neq B_1''(t_0). \quad (8)$$

Ruled surfaces are exactly in the first order contact at their common generator $a_{01}(t_0) = b_{01}(t_0)$.

From equalities of the dual arcs $ds_A = ds_B$ follows $H_A dt = H_B dt$, i.e. $H_A = H_B$. Since $B_1' = H_B \cdot B_2$, where $H_B = |B_1'|$, then from the condition $A_1'(t_0) = B_1'(t_0)$ we obtain that in common generator following equalities are holds:

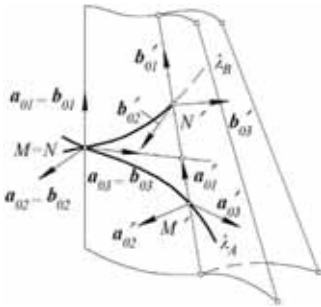


Figure 2: The n=1 order contact of the ruled surfaces

$$A_1 = B_1; A_2 = B_2; A_3 = B_3; H_A = H_B. \quad (9)$$

Proposition 1. For n = 1 order contact the ruled

surfaces have coincided trihedrons (Sannia frames) (a_{01}, a_{02}, a_{03}) and (b_{01}, b_{02}, b_{03}) at center point of their common generator (see Fig. 2).

1.3 The n=2 order contact of the ruled surfaces

$$A_1(t_0) = B_1(t_0); A_1'(t_0) = B_1'(t_0); \\ A_1''(t_0) = B_1''(t_0); A_1'''(t_0) \neq B_1'''(t_0). \quad (10)$$

Ruled surfaces are exactly in the second order contact at their common generator $a_{01}(t_0) = b_{01}(t_0)$ (see Fig. 3). From following relations:

$$A_1'' = (H_A \cdot A_2)' = \\ = -H_A^2 \cdot A_1 + H_A' \cdot A_2 + H_A \cdot Q_A \cdot A_3; \\ B_1'' = (H_B \cdot B_2)' = \\ = -H_B^2 \cdot B_1 + H_B' \cdot B_2 + H_B \cdot Q_B \cdot B_3,$$

it follows that additionally to results of n = 1 order contact, the following equalities are carried out

$$H_A' = H_B'; Q_A = Q_B. \quad (11)$$

Element of dual arc of ruled surface $A_3(t)$ on the basis of the derivation equation $A_3' = -Q_A \cdot A_2$ may be expressed as $ds_{A3} = |A_3'| \cdot dt = Q_A \cdot dt$. On the basis of the second equation of (11) it follows that $ds_{A3} = ds_{B3}$. Thus we obtain the equality of elements of dual arcs of ruled surfaces $A_3(t)$ and $B_3(t)$. And similarly on the basis of equations (11) and $A_2' = -H_A \cdot A_1 + Q_A \cdot A_3$ it follows that $ds_{A2} = ds_{B2}$. We obtain the equality of elements of dual arcs of ruled surfaces $A_2(t)$ and $B_2(t)$.

Proposition 2. For n = 2 order contact of ruled surfaces $A_1(t)$ and $B_1(t)$ the elements of dual arcs in pairs of surfaces $A_i(t)$ and $B_i(t)$; $i=1, 2, 3$, are equal and calculated from common generators in this pairs.

Equality of elements of dual arcs is equivalent to the equality of distribution parameters in the common generators of these pairs of surfaces. In theory of ruled surfaces [3] it is known that $\frac{ds_{A2}}{ds_A} = \frac{1}{\rho_A}$, where ρ_A - a dual curvature radius of ruled surface $A_1(t)$ in center point M of its ruling. As the $n = 2$ order contact is characterized by equality $ds_A = ds_B$ and $ds_{A2} = ds_{B2}$ it follows that in this contact ruled surfaces in their common generator $a_{01}(t_0) = b_{01}(t_0)$ have equal dual curvature radius $\rho_A = \rho_B$. Furthermore, since $\rho = \sin R$, where $R = R_0 + \omega R_1$ is a dual angle between the generator of ruled surface $A_1(t)$ and the generator of its first evolute (see Fig. 3) from equation $\rho_A = \rho_B$ it follows that $R_A = R_B$ and trihedrons of the first order evolutes are coincided.

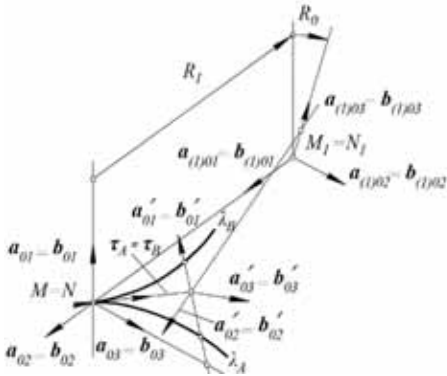


Figure 3: The $n=2$ order contact of the ruled surfaces

In the theory of complex line geometry [3] is known a formula for the three elements of dual arcs $ds_A = ds_{A1}$, ds_{A2} and ds_{A3} of ruled surfaces $A_1(t)$, $A_2(t)$ and $A_3(t)$ respectively:

$$ds_{A1}^2 + ds_{A3}^2 = ds_{A2}^2.$$

Since for $n=2$ order contact the following equalities are holds: $ds_{A1} = ds_{B1}$, $ds_{A2} = ds_{B2}$ and

$ds_{A3} = ds_{B3}$ so formula of three elements of arcs for ruled surfaces $B_1(t)$, $B_2(t)$ and $B_3(t)$ is holds:

$$ds_{B1}^2 + ds_{B3}^2 = ds_{B2}^2.$$

Obviously, the first order evolutes

$$A_{(1)3}(t) = a_{(1)03}(t) + \omega a_{(1)13}(t),$$

$$B_{(1)3}(t) = b_{(1)03}(t) + \omega b_{(1)13}(t), \quad \omega^2 = 0$$

of $n = 2$ order contact of ruled surfaces $A_1(t)$ and $B_1(t)$ are, in general, intersects along their common generator $a_{(1)03}(t_0) = b_{(1)03}(t_0)$.

The coincidence of trihedrons of evolutes at the center point $M_1 = N_1$ of their common generator is not sufficient for $n = 1$ order contact of these evolutes. As follows from the above studies of $n = 1$ order contact it requires equality of elements of dual arcs $ds_{A(1)3} = ds_{B(1)3}$ of first order evolutes $A_{(1)3}$ and $B_{(1)3}$.

Assume that the equality $ds_{A(1)3} = ds_{B(1)3}$ is hold. Then according to the formula of three elements of dual arcs for the ruled surface a dual equation is hold:

$$ds_{A(1)1}^2 + ds_{A(1)3}^2 = ds_{A(1)2}^2.$$

As $n = 2$ order contact the equality $ds_{A(1)1} = ds_{B(1)1}$ is holds and by assumption the equality $ds_{A(1)3} = ds_{B(1)3}$ is holds too, then from formula of three elements of dual arcs implies that $ds_{A(1)2} = ds_{B(1)2}$. Thus under our assumption, according with Proposition 2, we find that first order evolutes are in $n = 2$ order contact, that can't be true because in considering order contact of $A_1(t)$ and $B_1(t)$ only one equality $ds_{A(1)1} = ds_{B(1)1}$ is holds. Thus we have the inequality $ds_{A(1)3} \neq ds_{B(1)3}$ and first order evolutes $A_{(1)3}(t)$ and $B_{(1)3}(t)$ are intersecting along their common generator $a_{(1)03}(t_0) = b_{(1)03}(t_0)$.

Proposition 3. If ruled nondevelopable surfaces are in $n = 2$ order contact, their evolutes of the first order are intersects along their common generator.

Remark. Lets write elements of dual arcs of first order evolutes $A_{1(3)}(t)$ and $B_{1(3)}(t)$ in expanded form

$$\begin{aligned} ds_{A(1)3} &= ds_{0(A(1)3)} + \omega ds_{1(A(1)3)}, \\ ds_{B(1)3} &= ds_{0(B(1)3)} + \omega ds_{1(B(1)3)}, \quad \omega^2=0. \end{aligned}$$

Assume that $ds_{1(A(1)3)} = ds_{1(B(1)3)} = 0$. In this case the first order evolutes become developable surfaces and inequality $ds_{A(1)3} \neq ds_{B(1)3}$ results to real inequality $ds_{0(A(1)3)} \neq ds_{0(B(1)3)}$. According to the geometrical sense real values $ds_{0(A(1)3)}$ and $ds_{0(B(1)3)}$ are represents elements of arcs of developable evolute strictions [2] $A_{(1)3}(t)$ and $B_{(1)3}(t)$: $ds_{0(A(1)3)} = d\sigma_{A(1)3}$, $ds_{0(B(1)3)} = d\sigma_{B(1)3}$. According to the research of $n = 1$ order contact of developable surfaces [4] the following equality is holds: $k_{A(1)3} \cdot d\sigma_{A(1)3} = k_{B(1)3} \cdot d\sigma_{B(1)3}$, where $k_{A(1)3}$ and $k_{B(1)3}$ is strictions curvature of contacting surfaces. In general case this equality assumes the existing of inequality $d\sigma_{A(1)3} \neq d\sigma_{B(1)3}$. It is possible to assert that for some nondevelopable ruled surfaces which are in $n = 2$ order contact that there can be $n = 1$ order contact of their first order evolutes. This is possible only if evolutes are developable surfaces. This case is represented in Fig. 5.

1.4 The $n=3$ order contact of the ruled surfaces

$$\begin{aligned} A_1(t_0) &= B_1(t_0); A_1'(t_0) = B_1'(t_0); A_1''(t_0) = B_1''(t_0); \\ A_1'''(t_0) &= B_1'''(t_0); A_1^{IV}(t_0) \neq B_1^{IV}(t_0). \end{aligned} \quad (12)$$

Ruled surfaces are exactly in the third order contact at their common generator $a_{01}(t_0) = b_{01}(t_0)$. From following relations:

$$\begin{aligned} A_1''' &= -3H_A \cdot H'_A \cdot A_1 + \\ &+ (H''_A - H_A^3 - H_A \cdot Q_A^2) \cdot A_2 + \\ &+ (2H'_A \cdot Q_A + H_A \cdot Q'_A) \cdot A_3; \\ B_1''' &= -3H_B \cdot H'_B \cdot B_1 + \\ &+ (H''_B - H_B^3 - H_B \cdot Q_B^2) \cdot B_2 + \\ &+ (2H'_B \cdot Q_B + H_B \cdot Q'_B) \cdot B_3, \end{aligned}$$

it follows that additionally to results of $n = 2$ order contact the following equalities are carried out

$$H''_A = H''_B; Q'_A = Q'_B. \quad (13)$$

Thus in $n=3$ order contact following equations are holds:

$$\begin{aligned} \frac{ds_{A1}}{dt} &= H_A = H_B = \frac{ds_{B1}}{dt}; \\ \frac{d^2 s_{A1}}{dt^2} &= H'_A = H'_B = \frac{d^2 s_{B1}}{dt^2}; \\ \frac{d^3 s_{A1}}{dt^3} &= H''_A = H''_B = \frac{d^3 s_{B1}}{dt^3}; \\ \frac{ds_{A3}}{dt} &= Q_A = Q_B = \frac{ds_{B3}}{dt}; \\ \frac{d^2 s_{A3}}{dt^2} &= Q'_A = Q'_B = \frac{d^2 s_{B3}}{dt^2}. \end{aligned} \quad (14)$$

A center tangent's vector $A_3(t)$ has coordinates: $A_3(t) = \{x_3(t), y_3(t), z_3(t)\}$.

We define the derivatives of this vector minding the existence of functional relationship $s_{A1}(t)$:

$$\begin{aligned} A'_3 &= \frac{dA_3}{ds_{A1}} \cdot \frac{ds_{A1}}{dt} = \\ &= \left\{ \frac{dx_3}{ds_{A1}} \cdot \frac{ds_{A1}}{dt}; \frac{dy_3}{ds_{A1}} \cdot \frac{ds_{A1}}{dt}; \frac{dz_3}{ds_{A1}} \cdot \frac{ds_{A1}}{dt} \right\}. \end{aligned}$$

In the complex line geometry [3] are known equations:

$$\begin{aligned} \frac{dx_3}{ds_{A1}} &= -x_2 \cdot \text{ctg } R; \frac{dy_3}{ds_{A1}} = -y_2 \cdot \text{ctg } R; \\ \frac{dz_3}{ds_{A1}} &= -z_2 \cdot \text{ctg } R, \end{aligned} \quad (15)$$

where $A_2(t) = \{x_2(t), y_2(t), z_2(t)\}$.

We consider the getting a second derivative of the vector $A_3(t)$:

$$A''_3 = \left\{ \frac{d^2 x_3}{ds^2_{A1}} \cdot \frac{ds_{A1}}{dt} + \frac{dx_3}{ds_{A1}} \cdot \frac{d^2 s_{A1}}{dt^2}; \dots; \dots \right\}.$$

Proceeding from equations (15) it can be writ-

ten for the vector A_3' :

$$A_3'(t) = \left\{ \left(-\frac{x_{(1)2}}{\rho_{A1}} \operatorname{ctg} R + \frac{x_2}{\sin R} \cdot \frac{1}{r_{A1}} \right) \frac{ds_{A1}}{dt} + \right. \\ \left. + (-x_{(1)2} \operatorname{ctg} R) \frac{d^2 s_{A1}}{dt^2}; \dots; \dots \right\}, \quad (16)$$

where $A_{(1)2} = \{x_{(1)2}, y_{(1)2}, z_{(1)2}\}$, $\frac{1}{r_{A1}} = \frac{dR}{ds_{A1}}$,

ρ_{A1} and r_{A1} – are dual radius of curvature and torsion of a ruled surface $A_1(t)$ at generator of contacting.

Analysis of the components of vector $A_3''(t)$ in view of equations (14) and $dR + ds_{A(1)3} = 0$, which one is peculiar to evolute of ruled surface $A_1(t)$ [3], allows to draw certain conclusions. For $n = 3$ order contact of nondevelopable ruled surfaces the elements of dual arcs in pairs of surfaces $A_{(1)i}(t)$ and $B_{(1)i}(t)$, $i=1, 2, 3$ are equal.

Hence, the first order evolutes has $n=2$ order contact. In addition the second order evolutes trihedrons are coincide and the dual radius of torsion $r_{A1}=r_{B1}$ of $A_1(t)$ and $B_1(t)$ at parameter $t=t_0$ is equal.

2. STRICTION LINES BEHAVIOR

Consider the behavior of the strictions of contacting ruled surfaces. The striction's derivation equations [2] are of the form:

$$\frac{d\mathbf{x}_{(A)}}{dt} = q_{A1} \cdot \mathbf{a}_{01} + h_{A1} \cdot \mathbf{a}_{03}, \\ \frac{d\mathbf{x}_{(B)}}{dt} = q_{B1} \cdot \mathbf{b}_{01} + h_{B1} \cdot \mathbf{b}_{03}, \quad (17)$$

where $H_A = h_{A0} + \omega h_{A1}$, $Q_A = q_{A0} + \omega q_{A1}$,

$H_B = h_{B0} + \omega h_{B1}$, $Q_B = q_{B0} + \omega q_{B1}$, $\omega^2 = 0$.

For $n=1$ order contact on the basis of (9) we obtain:

$$\mathbf{a}_{01} = \mathbf{b}_{01}, \mathbf{a}_{02} = \mathbf{b}_{02}, \mathbf{a}_{03} = \mathbf{b}_{03}, \\ h_{A0} = h_{B0}, h_{A1} = h_{B1}. \quad (18)$$

Proposition 4: For $n=1$ order contact the strictions $\mathbf{x}_{(A)}$ and $\mathbf{x}_{(B)}$ of A_1 and B_1 are in-

tersects at central point M of common generator and their tangents are incident to tangent plane of surfaces at the same point.

For $n = 2$ order contact the conditions of (9) are added by conditions of (11). Then to equalities of (18) is added another equation

$q_{A0} + \omega q_{A1} = q_{B0} + \omega q_{B1}$, $\omega^2 = 0$ from which is follows $q_{A0} = q_{B0}$, $q_{A1} = q_{B1}$. In this case from (17) is follows that for $n = 2$ order contact the strictions $\mathbf{x}_{(A)}$ and $\mathbf{x}_{(B)}$ have common tangent vector at central point M of the common generator.

Proposition 5. For $n=2$ order contact the strictions $\mathbf{x}_{(A)}$ and $\mathbf{x}_{(B)}$ of $A_1(t)$ and $B_1(t)$ are contacting at central point of common generator. Their common tangent is incident to the tangent plane at central point.

Consider the behavior of the strictions $\mathbf{x}_{(A)}$ and $\mathbf{x}_{(B)}$ for $n=3$ order contact of $A_1(t)$ and $B_1(t)$. In the differential geometry is known that the curvature of the spatial curve can be defined by

$$k = \frac{|\llbracket \mathbf{x}', \mathbf{x}'' \rrbracket|}{|\mathbf{x}'|^3}. \quad (19)$$

Lets write the expressions of the numerator and denominator of the fraction (19) in relation to the striction $\mathbf{x}_{(A)}$ of $A_1(t)$ based on the first equation (17):

$$|\mathbf{x}'_{(A)}|^3 = (q_{A1}^2 + h_{A1}^2)^{3/2}. \quad (20)$$

$$\mathbf{x}''_{(A)} = q'_{A1} \cdot \mathbf{a}_{01} + (q_{A1} h_{A0} - h_{A1} q_{A0}) \mathbf{a}_{02} + \\ + h'_{A1} \cdot \mathbf{a}_{03}. \quad (21)$$

$$\llbracket \mathbf{x}'_{(A)}, \mathbf{x}''_{(A)} \rrbracket = -h_{A1} (q_{A1} h_{A0} - h_{A1} q_{A0}) \mathbf{a}_{01} + \\ + (q'_{A1} h_{A1} - h'_{A1} q_{A1}) \mathbf{a}_{02} + \\ + q_{A1} (q_{A1} h_{A0} - h_{A1} q_{A0}) \mathbf{a}_{03}. \quad (22)$$

Recall that for $n = 3$ order contact the equality $Q'_A = Q'_B$ is holds. By expanding this dual equality we obtain real equality: $q'_{A0} = q'_{B0}$, $q'_{A1} = q'_{B1}$. Furthermore, for $n = 2$ order contact the dual equality is carried out: $H'_A = H'_B$, from which is follows $h'_{A0} = h'_{B0}$, $h'_{A1} = h'_{B1}$. Applying the obtained

real equalities to the equations (20), (21), (22) and to the final formula (19) we arrive the result $k_{(A)}=k_{(B)}$, which means that for $n = 3$ order contact the strictions $x_{(A)}$ and $x_{(B)}$ at central point of the common generator $a_{01}=b_{01}$ have equal curvature. Further study of these strictions allows to found that their osculating planes at the central point of the common generator are coincide.

Proposition 6. In $n = 3$ order contact of ruled surfaces the strictions have the second order contact at a central point of common generator.

3. EXAMPLES

As an example we consider a contact of ruled surface $A_1 \rightarrow R_A(t, u)$ with elliptical hyperboloid $B_1 \rightarrow R_B(t, u)$ (see Fig. 4, 5):

$$A_1 \rightarrow R_A(t, u) = \left\{ A(\cos t + t \sin t) + \frac{u \cos t}{m}; \right.$$

$$\left. A(\sin t - t \cos t) + \frac{u \sin t}{m}; 1 + Bt + \frac{u}{m} \right\},$$

$$A = 1; B = 0.5; m = \sqrt{2}; 0 \leq t \leq 2\pi; -\infty < u < \infty.$$

$$B_1 \rightarrow R_B(t, u) = \left\{ c_1 \sin t + \frac{u \cos t}{\sqrt{1+c_3^2}}; \right.$$

$$\left. c_2 \cos t + \frac{u \sin t}{\sqrt{1+c_3^2}}; \frac{uc_3}{\sqrt{1+c_3^2}} \right\},$$

$$0 \leq t \leq 2\pi, -\infty < u < \infty.$$

$$n = 1: \quad c_1 = \frac{1-2t_0}{2c_3(3\cos^2 t_0 - 4)};$$

$$c_2 = \frac{-4(1-2t_0)}{2c_3(3\cos^2 t_0 - 4)}; c_3 = 2; t_0 = 2.46.$$

$$n = 2: \quad c_1 = \frac{2(1-t_0) - \cos^2 t_0 \cdot (3-4t_0)}{2(2\cos^2 t_0 - 1)};$$

$$c_2 = -\frac{1-2t_0 - \cos^2 t_0 \cdot (3-4t_0)}{2(2\cos^2 t_0 - 1)}; c_3 = 1; t_0 = 4.3$$

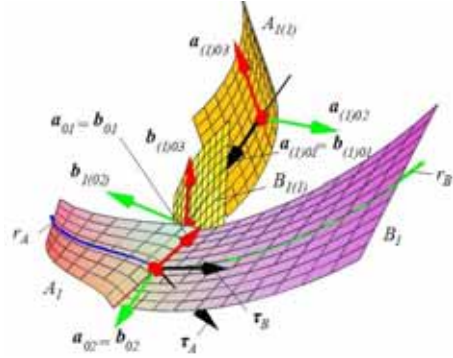


Figure 4: Example of $n=1$ order contact.

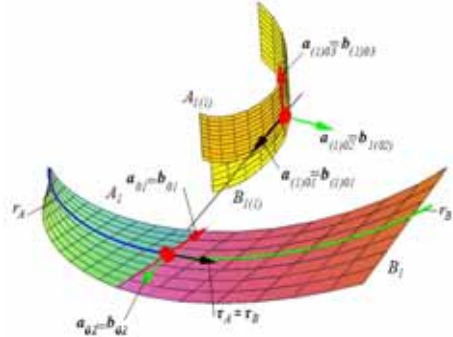


Figure 5: Example of $n=2$ order contact.

4. CONCLUSIONS

The n order contact of nondevelopable ruled surfaces is considered. The notion of a dual vector of divergence in common generator line of two contacting nondevelopable ruled surfaces is introduced via the dual vector calculus which describes the manifolds of line geometry. In meaning of this vector we prove necessary and sufficient conditions for the n order contact at the common generator line. Thus the primary orders of contact is researched. The behavior of the striction curves of contacting ruled surfaces is defined and some examples are considered. The results of these researches can be applied for design of composite ruled shells and in CAD systems via approximating issues.

REFERENCES

- [1] H. Pottmann and J. Wallner. Computational Line Geometry. Berlin, Springer Verlag, Heidelberg, 2001. –565 p.
- [2] Blaschke W. Vorlesungen über Differentialgeometrie. I Elementare Differentialgeometrie. 1 Aufl. Berlin, 1921, 230 s.
- [3] D. N. Zeiliger. Complex Line Geometry. Moscow-Leningrad, 1934, 196 p.
- [4] Elements of the theory of contact ruled developable surface. A.S. Niteyskiy, K.L. Panchuk. Journal of the Kuzbass State. tech. Univ., 2012, № 6(94). p. 112-117.

ABOUT THE AUTHORS

- 1. Konstantin Leonidovich PANCHUK, department of Engineer's Geometry & CAD, Omsk State Technical University, Russia.
email: Panchuk_KL@mail.ru
- 2. Anton Sergeevich NITEYSKIY, department of Engineer's Geometry & CAD, Omsk State Technical University, Russia.
email: AntonGTH@gmail.com

CONTROL METHOD OF COMBINED DEVELOPABLE SURFACE DESIGN BY AFFINE TRANSFORMATION AND LOCUS

Hiroataka SUZUKI
Kobe University, Japan

ABSTRACT: Developable surfaces have significant advantage in manufacturing process, though designing of the surfaces have much restrictions. Author have already proposed combined developable surface design method and introduced the method to CG exercise class of graphic science education. In this method, each developable surface which composes designed shape is generated by manipulation (translation, scaling and rotation) of curved line. Though this method have much design potential, complexity level of designed examples is remaining in a certain level. Because if manipulation of curved line is extended to much complicated level, the features of designed shapes cannot be easily understood at designing phase. To solve this problem, author introduced affine transformation expression for understanding features of designed shapes and locus diagram for visualization of series of manipulation. In this paper, principle of proposed design method is explained simply at first. Then relationship among manipulation of curved line, affine transformation and locus diagram is explained. Next, examples of affine transformation expression and locus diagram expression of designed shapes by manipulation in higher level are indicated. Finally, submitted works by students are shown as examples of proposed design process.

Keywords: Developable surface, Cylindrical surface, Conical surface, Directing line, Generating line.

1. INTRODUCTION

Developable surfaces have significant advantage in manufacturing process, though designing of the surfaces have much restrictions. Therefore a lot of methods have been proposed to combine developable surfaces to increase freedom of shape design. The research about developable surfaces are classified into two categories, research which enable approximation of given curved surface with combination of developable surfaces and research which enable easy designing with combination of developable surfaces.

As a research of easy designing, author have already proposed combined developable surface design method[1] and introduced the method to CG exercise class of graphic science education[2]. As shown in Figure 1, each developable surface which composes designed shape is generated by manipulation (translation, scaling and rotation) of curved line in this

method. Though this method have much design potential, complexity level of designed examples is remaining in a certain level. Because if manipulation of curved line is extended to much complicated level, for example manipulation which includes combination of rotations around different axes as shown in Figure 2, the features of designed shape cannot be easily

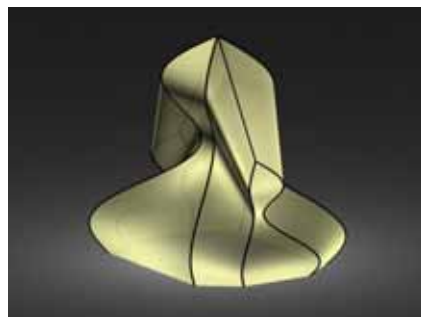


Figure 1: An example of designed shape generated by manipulation of given curve [1].

understood at designing phase.

To solve this problem, author introduced affine transformation for understanding features of designed shape and locus diagram for visualization of series of manipulation. As each manipulation in proposed design method is corresponding to affine transformation mathematically, combination of manipulations can be expressed as product of affine transformation matrix. With expression of product of matrix, features of designed shape, for examples property of continuity or symmetry of the shape, can be easily understood. And each manipulation of proposed design method can be expressed by locus diagram which can easily express content of manipulations visually and simply.

In this paper, principle of proposed design method is explained simply at first. Then relationship among manipulation of curved line, affine transformation and locus is explained. Next, examples of affine transformation expression and locus diagram expression of designed shapes by manipulation in higher level are indicated. Finally, submitted works by students are shown as examples of proposed design process.

Iannis Xenakis, an architect, once tried similar designing method. When he designed 'Philips Pavilion (1958, Brussels)', he combined hyperbolic paraboloid surfaces and conical surfaces as shown in Figure 3. In proposing method, some of directing lines are become intersection lines between combined surfaces. Different from proposing method, some of generating lines are become intersection lines between combined surfaces in method by Xenakis.

Some of globes on the market are approximation of sphere making use of several cylindrical surfaces. Or, Mitani[3] proposed manufacturing method of approximated solid of revolution by folded paper (ORIGAMI). These methods can be thought as a part of proposing method. Proposing method is further developed and generalized comparing to these existing methods.

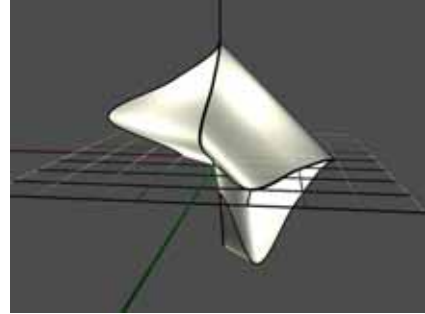


Figure 2: An example of designed shape generated by manipulation which includes combination of rotations around different axes.

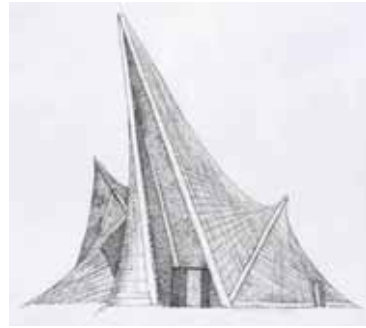


Figure 3: A sketch of Philips Pavilion (1958, Brussels) drawn by Prof. Koji MIYAZAKI.

Though designing of developable surfaces have much restrictions, manufacturing of developable surfaces is very easy. Therefore a lot of methods have been proposed to combine developable surfaces to increase freedom of shape design. These methods are classified to 2 categories, research which enable approximation of given curved surface with combination of developable surfaces and research which enable easy designing with combination of developable surfaces.

As research for approximation, Mitani et al. [4] proposed the method of approximation with plane triangles, Shatz et al.[5] proposed the method of approximation with conical surfaces and planes, Massawri et al.[6] proposed the

method of approximation with tubes which are constructed by triangles and Pottman et al.[7] proposed the method of approximation with developable strips.

As research for easy designing, Rose et al.[8] proposed the method of 3 dimensional shape generation from 2 dimensional perspective drawing, Kilian et al.[9] proposed method of shape generation with repetition of curved line folding and Suzuki[10] proposed method of shape generation with combination of tangent surfaces.

The objective of the proposing method in this paper is offering the way of designing and manufacturing for graphic science education and design workshop of restricted time. The method belongs to the research for easy designing. Though proposed method restricts freedom of design in certain level, the method enables designing and manufacturing in very short time.

In this paper, right-handed coordinate system as shown in Figure 4 is adopted for coordinate expression.

2. EXPRESSION OF CURVED LINE MANIPULATION BY AFFINE TRANSFORMATION

2.1 Relations between manipulation and affine transformation

In general, affine transformation is matrix transformation expressed by following Equation (1):

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} a_{41} \\ a_{42} \\ a_{43} \end{pmatrix} \quad (1)$$

Considering easy operation for combination of transformations, Equation (1) is often expressed as homogeneous coordinate system as following Equation (2):

$$\begin{pmatrix} x' \\ y' \\ z' \\ 1 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{21} & a_{31} & a_{41} \\ a_{12} & a_{22} & a_{32} & a_{42} \\ a_{13} & a_{23} & a_{33} & a_{43} \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} \quad (2)$$

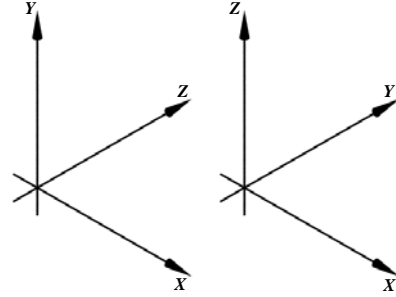


Figure 4: Left-handed coordinate system(left) and right-handed coordinate system(right).

In this paper, affine transformation is expressed as homogeneous coordinate system as in Equation (2). If we assume transformation for translation with moving vector (x_t, y_t, z_t) as $\mathbf{T}(x_t, y_t, z_t)$, transformation for scaling from origin point with scaling rate k ($0 < k < 1$) as $\mathbf{S}(k)$ and transformation for rotation around XYZ axis with rotating angle θ as $\mathbf{R}_x(\theta)$, $\mathbf{R}_y(\theta)$, $\mathbf{R}_z(\theta)$, these transformations are expressed as following matrix:

$$\mathbf{T}(x_t, y_t, z_t) = \begin{pmatrix} 1 & 0 & 0 & x_t \\ 0 & 1 & 0 & y_t \\ 0 & 0 & 1 & z_t \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (3)$$

$$\mathbf{S}(k) = \begin{pmatrix} k & 0 & 0 & 0 \\ 0 & k & 0 & 0 \\ 0 & 0 & k & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (4)$$

$$\mathbf{R}_x(\theta) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (5)$$

$$\mathbf{R}_y(\theta) = \begin{pmatrix} \cos\theta & 0 & \sin\theta & 0 \\ 0 & 1 & 0 & 0 \\ -\sin\theta & 0 & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (6)$$

$$\mathbf{R}_z(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 & 0 \\ \sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (7)$$

Table 1: Implemented functions in POV-Ray and content of the functions.

Functions in POV-Ray	Content of function
Translate(<xt, yt, zt>)	$T(xt, yt, zt)$ Equation (3)
Scale(k)	$S(k)$ Equation (4)
Rotate_X(theta)	$R_x(\theta)$ Equation (5)
Rotate_Y(theta)	$R_y(\theta)$ Equation (6)
Rotate_Z(theta)	$R_z(\theta)$ Equation (7)
Scale_frompoint(k, <xs, ys, zs>)	$S'(k, xs, ys, zs)$ Equation(8)
Rotate_Xparallel(theta, <xr, yr, zr>)	$R'_x(\theta, xr, yr, zr)$ Equation (9)
Rotate_Yparallel(theta, <xr, yr, zr>)	$R'_y(\theta, xr, yr, zr)$ Equation (10)
Rotate_Zparallel(theta, <xr, yr, zr>)	$R'_z(\theta, xr, yr, zr)$ Equation (11)

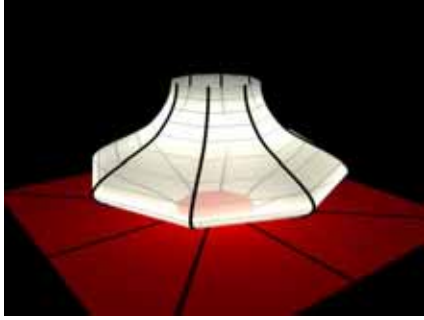


Figure 5: An example shape made from 8 times rotations of given curve.

And, if we assume transformation of scaling from non-origin point (x_s, y_s, z_s) with scaling rate k as $S'(k, x_s, y_s, z_s)$ and transformation of rotation around the line parallel to XYZ axis with rotating angle θ as $R'_x(\theta)$, $R'_y(\theta)$, $R'_z(\theta)$, these transformations are expressed by combination of matrix in Equation (3) to (7) as following:

$$S'(k, x_s, y_s, z_s) = T(x_s, y_s, z_s) S(k) T(-x_s, -y_s, -z_s) \quad (8)$$

$$R'_x(\theta, x_r, y_r, z_r) = T(x_r, y_r, z_r) R_x(\theta) T(-x_r, -y_r, -z_r) \quad (9)$$

$$R'_y(\theta, x_r, y_r, z_r) = T(x_r, y_r, z_r) R_y(\theta) T(-x_r, -y_r, -z_r) \quad (10)$$

$$R'_z(\theta, x_r, y_r, z_r) = T(x_r, y_r, z_r) R_z(\theta) T(-x_r, -y_r, -z_r) \quad (11)$$

2.2 Description of transformation matrix in CG freeware POV-Ray

POV-Ray is freeware application for CG modeling and rendering. All objects in a scene are located by text expression in POV-Ray. As same as description of general computer

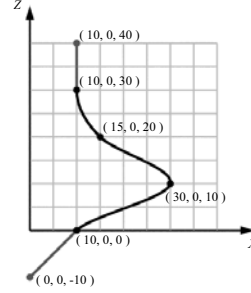


Figure 6: Six control points and a spline curved line generated from the points.

```
#declare Number_of_ControlPoints = 6;
#declare Points_for_CurvedLine
= array[ MAX_NUMBER_POINT ][ 2 ]
{
    { 0, -10 }, // point 1
    { 10, 0 }, // point 2
    { 30, 10 }, // point 3
    { 15, 20 }, // point 4
    { 10, 30 }, // point 5
    { 10, 40 }, // point 6
    // point 7 and after
    { 0, 0 }, { 0, 0 }, { 0, 0 }, { 0, 0 }, { 0, 0 }, { 0, 0 }, { 0, 0 },
    { 0, 0 }, { 0, 0 }, { 0, 0 }, { 0, 0 }, { 0, 0 }, { 0, 0 }, { 0, 0 }
}
```

Figure 7: A definition of spline curved line for the shape shown in Figure 5.

program, repetition or conditional branch or functions can be used in POV-Ray. As shown in Table 1, we implemented all transformation matrices shown in Equation (3) to (11) as function in POV-Ray.

2.3 Description of transformation matrix in CG freeware POV-Ray

Figure 5 shows an example shape made from 8 times rotations of given curve. Rotation angles in 8 rotations are the same angle $\pi/4$. Given curve is spline curve defined from six control points as shown in Figure 6. In six control points, the first and the last points are defined for angle at edges of the curved line. As shown in Figure 7 and 8, definitions for spline curve (number of control points and coordinates of control points) and definition of manipulations (number of manipulations and content of ma-

```

#declare Number_of_Manipulations = 8;
#declare A[1] = Rotate_Z(2*pi/8);
#declare A[2] = Rotate_Z(2*pi/8);
#declare A[3] = Rotate_Z(2*pi/8);
#declare A[4] = Rotate_Z(2*pi/8);
#declare A[5] = Rotate_Z(2*pi/8);
#declare A[6] = Rotate_Z(2*pi/8);
#declare A[7] = Rotate_Z(2*pi/8);
#declare A[8] = Rotate_Z(2*pi/8);

```

Figure 8: A definition of manipulations for the shape shown in Figure 5.

```

object {
  combination_of_developablesurfaces (
    Number_of_Manipulation,
    0, //Draw all surfaces if the number is 0
    //Draw directed surface if the number is not 0
    Number_of_ControlPoints,
    Points_for_CurvedLine,
    A //Affine transformations
  )
}

```

Figure 9: An example of the function which constructs the shape from given curved line and given transformations.

nipulations) are declared beforehand. Then these definitions are passed to shape construction function 'combination_of_developablesurfaces' as shown in Figure 9. After camera settings (location, direction and angle of view) and information of other objects are described, user can obtain perspective view of defined shape.

As shown in Figure 7, curved line is defined as plane curve. When given manipulations do not include rotation, space curve can be used to get developable surface. At present, only plane curve can be given in this system.

Figure 10 is another example of the shape made from manipulation of given curved line. In this shape, 12 manipulations including translation, scaling and rotation are used as shown in Figure 11.

2.4 Understanding of features of shapes with affine transformation

The manipulations given for constructing the shape shown in Figure 5 can be expressed as following equation (12):



Figure 10: An example shape made from 12 manipulations of given curve.

```

#declare A[1] = Scale_frompoint(3/4, <40, 40, 0>);
#declare A[2] = Rotate_Zparallel(pi/2, <10, 10, 0>);
#declare A[3] = Scale_frompoint(4/3, <40, 40, 0>);

#declare p_offset = <-5, -5, 0>;

#declare A[4] = Translate(p_offset);
#declare A[5] = Rotate_Zparallel(pi*4, p_offset);
#declare A[6] = Rotate_Zparallel(pi*4, p_offset);
#declare A[7] = Scale_frompoint(3/4, <-40, -40, 0> +
  p_offset);
#declare A[8] = Rotate_Zparallel(pi/2, <-10, -10, 0> +
  p_offset);
#declare A[9] = Scale_frompoint(4/3, <-40, -40, 0> +
  p_offset);
#declare A[10] = Rotate_Zparallel(pi/4, p_offset);
#declare A[11] = Rotate_Zparallel(pi/4, p_offset);
#declare A[12] = Translate(-p_offset);

```

Figure 11: A definition of manipulations for the shape shown in Figure 10.

$$\mathbf{A}^8 \quad (\text{where } \mathbf{A} = \mathbf{R}_z(2\pi/8)) \quad (12)$$

Obviously, \mathbf{A}^8 equals unit matrix. If given curved line overlap to original line after a series manipulations, the product of all matrices corresponding to manipulations must be unit matrix. In this case, the last curved surface is connected to the first without excess and deficiency and users can easily understand this feature strictly from the viewpoint of mathematics.

In case of the shape shown in Figure 10, manipulations for given curved line are described in Figure 11 and the manipulations can be described in affine transformation format as following equation:

$$\begin{aligned}
& T(-P)R'_z\left(\frac{\pi}{4}, P\right)R'_z\left(\frac{\pi}{4}, P\right) \\
& \times S'\left(\frac{4}{3}, -45, -45, 0\right)R'_z\left(\frac{\pi}{2}, -15, -15, 0\right) \\
& \times S'\left(\frac{3}{4}, -45, -45, 0\right)R'_z\left(\frac{\pi}{4}, P\right)R'_z\left(\frac{\pi}{4}, P\right)T(P) \\
& \times S'\left(\frac{4}{3}, 40, 40, 0\right)R'_z\left(\frac{\pi}{2}, 10, 10, 0\right)S'\left(\frac{3}{4}, 40, 40, 0\right) \quad (13)
\end{aligned}$$

where $P = (5, -5, 0)$.

The product of 12 matrices in Equation (13) is actually unit matrix and as same as the shape shown in Figure 5, the last curved surface is connected to the first without excess and deficiency. And, $A[3]$ and $A[1]$, $A[4]$ and $A[12]$, $A[5]$ and $A[11]$, $A[6]$ and $A[10]$ are symmetry with respect to the surface $y = x$. And each transformation of $A[2]$ and $A[8]$ are also symmetry with respect to the surface $y = x$. We can understand that the shape constructed by transformations described in Figure 11 or Equation (13) is symmetry with respect to the surface $y = x$.

In this way, users can understand features of constructed shape, symmetry or continuity, by considering manipulations as affine transformations.

3. EXPRESSION OF CURVED LINE MANIPULATION BY LOCUS DIAGRAM

3.1 Locus diagram expression of translation, scaling and rotation of curved line.

A point given in spatial field may move by manipulation matrices. Locus diagram expression of manipulation enables easy understanding for series of manipulations. To realize easy understanding of manipulations, we developed function of generating locus of given 2 points, main point and sub point as shown in Figure 12 to 14. Though locus diagram can be drawn with only 1 point, 2 points are needed to make sure direction of manipulated curved line.

In case of rotation, given point may move along an arc of circle. However, generating lines on the cylindrical surface generated by rotation of plane curve are straight lines in this design method. Considering the difference, locus lines by rotation are drawn as straight lines

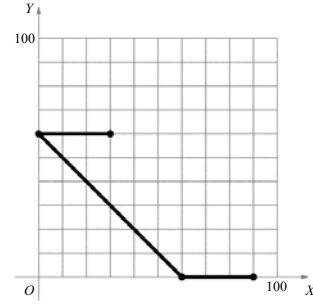


Figure 12: An example of locus diagram corresponding to translation $T(-60, 60, 0)$.

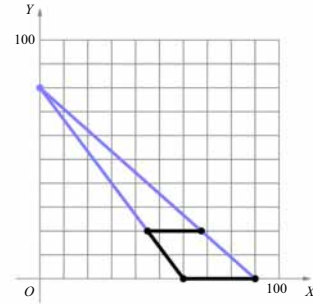


Figure 13: An example of locus diagram corresponding to scaling $S'(3/4, 0, 80, 0)$.

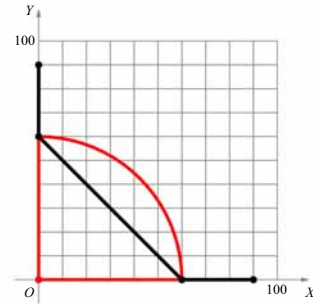


Figure 14: An example of locus diagram corresponding to rotation $R_z(\pi/2)$.

with arcs which indicate the given manipulations are rotation as shown in Figure 14.

Concretely, in case of translation, the segment which links main point and sub point, and the line which corresponding to manipulated segment are drawn in black. And the segment

which links main point and manipulated main point is drawn also as shown in Figure 12. In case of scaling, 3 black lines are drawn as same as case of translation. In addition to these 3 black lines, 2 blue lines are also drawn among center point of scaling, main point and sub point as shown in Figure 13. In case of rotation, 3 black lines are drawn as same as case of translation also. In addition to these lines, 2 red lines are also drawn among center point of rotation, main point and manipulated main point, and the arc which corresponding to manipulation of the rotation is drawn as shown in Figure 14. We can easily discriminate manipulations in Figure 12 from manipulation in Figure 14 though start point and manipulated start point are exactly same in this case.

Though an original idea of locus diagram was indicated by Suzuki[1], several information is added to realize easy understanding of manipulations in proposed diagram. Mitani[11] also proposed a design method of folded paper with given profile polyline and given trajectory polyline. There are similar points between proposed locus diagram and design method by Mitani. In case of method by Mitani, there is advantageous point that 2 developable surfaces are generated by folding along given polyline, though trajectory polyline do not include scaling which is included in proposing method.

3.2 Understanding of features of shapes with locus diagram

Figure 15 shows the locus diagram corresponding to series of manipulations shown in Figure 5. The shape in the figure is obviously rotational symmetry and continuous without excess and deficiency. Though, the features can be understood from affine transformation, the locus diagram shown in Figure 15 can tell the features easier. Figure 16 shows the locus diagram corresponding to series of manipulations shown in Figure 10. Though continuity and symmetry of the shape are mentioned in 2.4 from viewpoint of affine transformation, we can understand the features easier with the locus diagram. Strict features cannot be under-

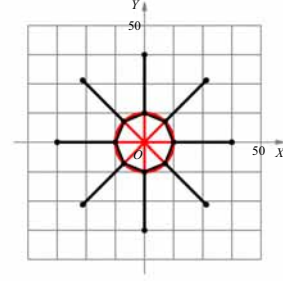


Figure 15: The locus diagram corresponding to series of manipulations shown in Figure 5.

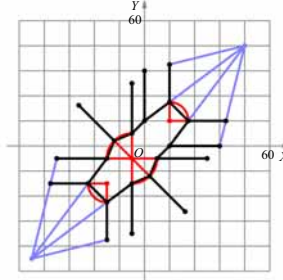


Figure 16: The locus diagram corresponding to series of manipulations shown in Figure 10.

stood only with locus diagram, however, we can understand the features intuitively and strictly with both affine transformation expressions and locus diagram expressions.

4. EXAMPLES OF DESIGNED SHAPES AND EXPRESSIONS OF THE SHAPES

In this chapter, several examples of designed shapes and affine transformation expressions and locus diagram expressions of corresponding shapes are indicated.

At first, an example of designed shape made from series of manipulations given as Equation (14) is shown.

$$\begin{aligned} & \mathbf{R}'_z\left(\frac{\pi}{4}, p_3\right) \mathbf{R}'_z\left(\frac{\pi}{4}, p_3\right) \\ & \times \mathbf{R}'_z\left(-\frac{\pi}{8}, p_2\right) \mathbf{R}'_z\left(-\frac{\pi}{8}, p_2\right) \mathbf{R}'_z\left(-\frac{\pi}{8}, p_2\right) \mathbf{R}'_z\left(-\frac{\pi}{8}, p_2\right) \\ & \times \mathbf{R}'_z\left(\frac{\pi}{4}, p_1\right) \mathbf{R}'_z\left(\frac{\pi}{4}, p_1\right) \end{aligned} \quad (14)$$

where $l = 24$, $p_1 = (l, 0, 0)$, $p_2 = (l, l, 0)$

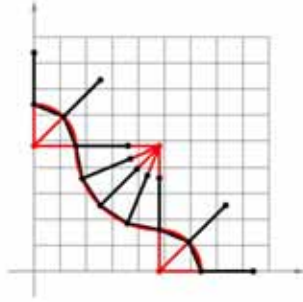


Figure 17: The locus diagram corresponding to manipulations shown in Equation (14).

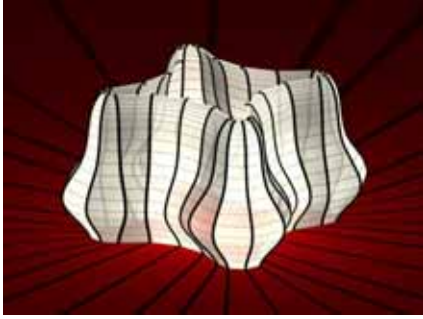


Figure 18: The shape generated from manipulations shown in Equation (14) 4 times.

$$\text{and } p_3 = (0, l, 0)$$

Figure 17 shows the locus diagram expression of given manipulations in Equation (14), and Figure 18 shows the shape generated by giving the manipulations 4 times. So far, in this paper, shapes generated from rotations of same direction are introduced. As shown in Figure 18, rotations of inverse directions can be given to original curved line.

Next, Figure 19 shows the locus diagram of 12 manipulations expressed as Equation (15)

$$\begin{aligned} & \mathbf{S}'\left(\frac{95}{100}, 0, 95, 0\right) \mathbf{R}'_z\left(\frac{\pi}{4}, 0, -5, 0\right) \mathbf{R}'_z\left(\frac{\pi}{4}, 0, -5, 0\right) \\ & \times \mathbf{S}'\left(\frac{95}{100}, 95, -5, 0\right) \mathbf{R}'_z\left(\frac{\pi}{4}, -5, -5, 0\right) \mathbf{R}'_z\left(\frac{\pi}{4}, -5, -5, 0\right) \\ & \times \mathbf{S}'\left(\frac{95}{100}, -5, -100, 0\right) \mathbf{R}'_z\left(\frac{\pi}{4}, -5, 0, 0\right) \mathbf{R}'_z\left(\frac{\pi}{4}, -5, 0, 0\right) \\ & \times \mathbf{S}'\left(\frac{95}{100}, -100, 0, 0\right) \mathbf{R}'_z\left(\frac{\pi}{4}, 0, 0, 0\right) \mathbf{R}'_z\left(\frac{\pi}{4}, 0, 0, 0\right) \quad (15) \end{aligned}$$

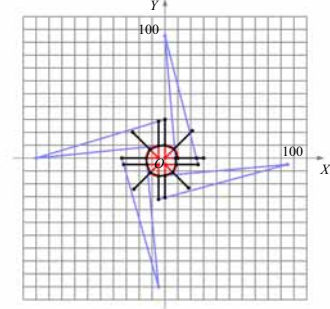


Figure 19: The locus diagram corresponding to manipulations shown in Equation (15).

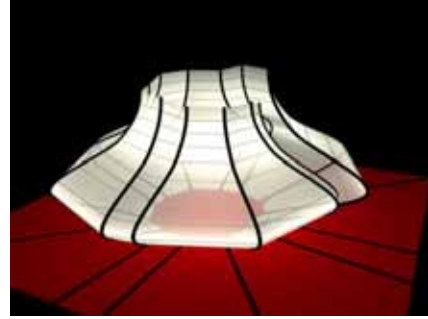


Figure 20: The shape generated from manipulations shown in Equation (15) and Figure 19.

So far, in this paper, shapes in which the last curved surface is connected to the first without excess and deficiency are introduced. As shown in Figure 19, spiral shape in which the last curved surface is not connected to the first can be generated by combination of manipulations. The shape shown in Figure 20 is an example which is made from manipulations described as Equation (15). Original curved line for the shape shown in Figure 20 is exactly same as that for the shape shown in Figure 5. Though given manipulations for the shape in Figure 20 is similar to that for the shape in Figure 5, scaling manipulations generate such spiral

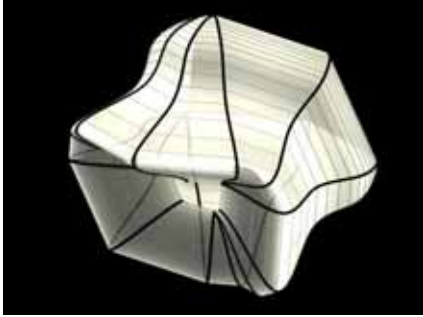


Figure 21: The shape generated from manipulations shown in Equation (16) and Figure 19.

shape.

At last, Figure 21 shows the shape generated from series manipulations described as Equation (16).

$$\begin{aligned} &R_x(-\pi/4)R_x(-\pi/4)R_y(-\pi/4)R_y(-\pi/4)R_z(\pi/4)R_z(\pi/4) \\ &\times R_x(-\pi/4)R_x(-\pi/4)R_y(-\pi/4)R_y(-\pi/4)R_z(\pi/4)R_z(\pi/4) \quad (16) \end{aligned}$$

So far, in this paper, shapes generated from rotations around lines which are parallel to Z axis are introduced. As shown in the figure, a series of manipulations can include rotations in which the axes of the rotations are not parallel to each

other. Though it is very difficult to consider continuity of the shape generated from such manipulations from perspective view of the shape, we can understand continuity strictly from affine transformation matrices.

5. IMPLEMENTATION OF THE DESIGN PROCESS IN GRAPHIC SCIENCE EDUCATION COURSE

We tried to implement the design process explained at this paper in the graphic science education course. We implemented the process in the class titled 'Environmental Equipment Planning' which is provided for 1st grade graduate students of department of architecture, graduate school of engineering, Kobe university. In academic year 2014, 4 units (90 minutes / unit) were allocated for lampshade design exercise. Students were requested to design lampshade with the design process and manufacture the lampshade actually. In total, 18 works were submitted. Figure 22 shows the examples of submitted works by CG perspective view, locus diagram and photo of manufactured works.

6. CONCLUSIONS AND FUTURE WORK

To improve usability of design method which combine developable surfaces with given curved line and given manipulations, and to

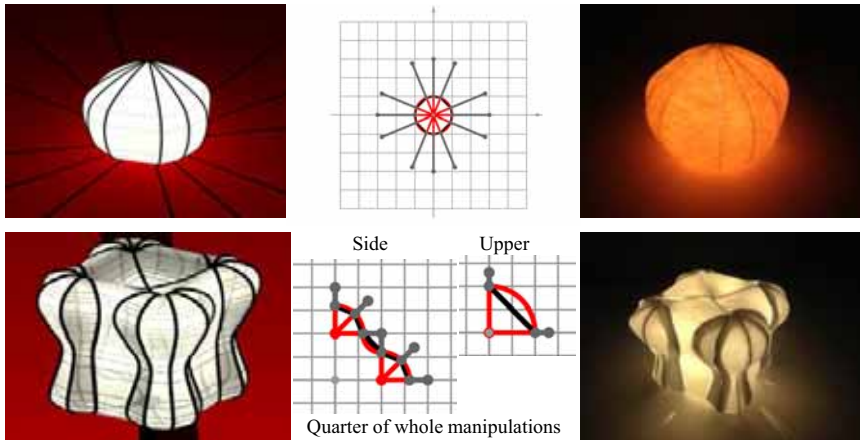


Figure 22: The examples of submitted works by CG perspective view, locus diagram and photo of manufactured works

improve understanding of features of generated shapes, expressions of manipulations by affine transformation and by locus diagram are proposed. With expression of affine transformation, we can strictly understand the features in detail, and with locus diagram we can intuitively and roughly understand the features. We are planning to add mirroring as one of manipulation. If we adopt mirroring, generated curved surface from space curve with mirroring is cylindrical surface. As shown in Figure 7, only plane curve is allowed as given curved line at current system. We have to extend the system to allow space curve as given curved line to make good use of advantageous point by introduction of mirroring.

ACKNOWLEDGMENTS

Examples of works shown in chapter 5 were submitted by graduate course students of department of architecture, graduate school of engineering, Kobe university. Author would like to thank all students who participated lampshade design exercise and submit their own lampshade within a given period of time.

REFERENCES

- [1] H. Suzuki, A generation method of developable surface by manipulation (translation, scaling and rotation) of curved line and a designing method of complicated shapes by combination of generated developable surfaces, *Journal of Graphic Science of Japan*, in Japanese, 48(1), pp.3-10, 2014.
- [2] H. Suzuki, Ai Sakaki, Kensuke Yasufuku and Takashi Matsumoto, Designing of Lampshade with 3D CG Application and Manufacturing of Designed Shape in Graphic Science Education, *Proceedings of The 2013 Asian Forum on Graphic Science*, pp.38-43, 2013
- [3] J. Mitani, A Design Method for 3D Origami Based on Rotational Sweep, *Computer-Aided Design and Applications*, 6(1), 69-79, 2009.
- [4] J. Mitani, Hiromasa Suzuki, Making Papercraft Toys from Meshes using Strip-based Approximate Unfolding, *ACM Trans., Graphics*, 23(3), pp.259-263, 2004.
- [5] I. Shatz, A. Tal, G. Leifman, Paper craft models from meshes, *The Visual Computer: International Journal of Computer Graphics archive*, Vol. 22 Issue 9, pp.825-834, 2006.
- [6] F. Massarwi, C. Gotsman and G. Elber, Papercraft Models Using Generalized Cylinders, *Proceedings of Pacific Graphics*, pp.148 – 157, 2007.
- [7] H. Pottmann, A. Schiftner, P. Bo, H. Schmiedhofer, W. Wang, N. Baldassini and J. Wallner, Freeform surfaces from single curved panels, *ACM Trans. Graphics*, 27(3), Article No.76, 2008.
- [8] K. Rose, A. Sheffer, J. Wither, M. Cani and B. Thibert, Developable Surfaces from Arbitrary Sketched Boundaries, *Eurographics Symposium on Geometry Processing*, pp.163-172, 2007.
- [9] M. Kilian, S. Flory, Z. Chen, N. J. Mitra, A. Sheffer and H. Pottmann, Curved Folding, *ACM Trans. Graphics*, 27(3), Article No.75, 2008.
- [10] H. Suzuki, Designing of Lighting Equipment making use of Tangent Surface and Control Method of the Surface by Hermite Curve, *Proc. of 15th Intl. Conference on Geometry and Graphics*, in CDROM, 2012.
- [11] J. Mitani, Column-shaped Origami Design Based on Mirror Reflections, *Journal for Geometry and Graphics*, Vol.16, No.2, pp.185-194, 2012.
- [12] H. Suzuki, A Study on Impact of Introduction of Lighting Equipment Design Assignment into Graphic Science Education, *Proc. of 12th Intl. Conference on Geometry and Graphics*, in CDROM, 2006.
- [13] H. Suzuki, Application of tangent surfaces in the design of lampshades, *Proc. of the Fall Annual Conference of the Japan Society for Graphics Science*, 133-136, in Japanese, 2010.

- [14]H. Suzuki, 'Education Course of Light and Shape Designing making use of Folded Papers', Proc. of the Fall Annul Conference of the Japan Society for Graphics Science, 143-144, in Japanese, 2011.
- [15]H. Suzuki, 'Designing of Lighting Equipment Making Use of Tangent Surface and Control Method of the Surface by Hermite Curve', Proc. of 15th Intl. Conference on Geometry and Graphics, in CDROM, 2012.

ABOUT THE AUTHORS

1. Hirotaka SUZUKI, Dr. Eng., is an associate professor of Department of Architecture, Graduate School of Engineering, Kobe University. His research interests are Lighting Environment Simulation, Geometrical Design and Graphic Science Education. He can be reached by e-mail: hirotakasuzuki@people.kobe-u.ac.jp, or through postal address: 1-1, Rokkodai-cho, Nada-ku, Kobe City, 657-8501, JAPAN.

CURVATURE FUNCTIONS ON A ONE-SHEETED HYPERBOLOID

Boris ODEHNAL

University of Applied Arts Vienna, Austria

ABSTRACT: We study the distribution of some curvature functions on a one-sheeted hyperboloid by determining, describing, and visualizing the curves of constant Gaussian, Mean, principal curvature, and the curves of constant ratio of the principal curvatures. Our aim is a precise description of the regions of prescribed curvature values. It turns out that all these curves are algebraic and can be given in terms of implicit equations. Surprisingly, it is possible to derive an explicit parametrization of the curves of constant principal curvature in terms of algebraic functions.

Keywords: One-sheeted hyperboloid, Gauss curvature, Mean curvature, principal curvature, constant curvature, support function, striction curve, ratio of principal curvatures, principal view.

1. INTRODUCTION

Some CAD-systems offer tools for *curvature analysis*. With these tools parts of surfaces can be textured with a color map of the Gaussian and Mean curvature. The primitives in a CAD system are usually approximated by some free-form surfaces. Thus, the curvature analysis sometimes tends to produce strange results. Symmetries of surfaces cause symmetries in the distribution of curvatures (cf. Figure 1). Unfortunately, these symmetries are not reproduced by the curvature analysis tools. We aim at a precise description of the distribution of some well-known curvature functions on a one-sheeted hyperboloid. The case of a hyperboloid of revolution is not treated here, since almost all of the curves we are dealing with are parallel circles in this case. The respective curves on an ellipsoid are studied in [6].

In Section 2 we study the distribution of the Gauss curvature. Then, Section 3 is devoted to the curves of constant Mean curvature. Finally, in Section 4 we derive and investigate the curves of constant principal curvature together with the curves of constant ratio of the two principal curvatures. We derive an explicit parametrization of the curves of constant principal curvature in terms of algebraic functions. Note that these curves do not agree with the principal curvature lines.

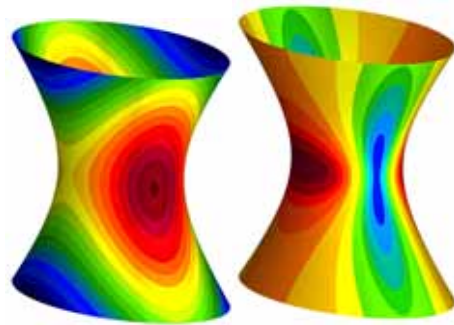


Figure 1: Gaussian curvature (left) and mean curvature (right) on a one-sheeted hyperboloid: Regions of a certain color correspond to curvature values within some interval.

In the following, when we describe the iso-curves of some curvature functions, we use the triplet of orthogonal projections onto the three mutually orthogonal planes of symmetry of the quadric. These planes shall coincide with the coordinate planes and we call the images appearing in the $[x, y]$ -, $[y, z]$ -, and $[x, z]$ -plane the *top view*, the *front view*, and the *(right) side view*.

2. CONSTANT GAUSSIAN CURVATURE

The surface \mathcal{S} in question shall be the one-sheeted hyperboloid with the equation

$$\mathcal{S} : \frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1 \quad (1)$$

where $0 < a < b$ and $c > 0$ are real coefficients. It is well-known that \mathcal{S} can be considered as a ruled surface in two different ways: Both ruled surfaces

$$R_{1,2}(u, v) = \begin{pmatrix} a \cos u \\ b \sin u \\ 0 \end{pmatrix} + v \cdot \begin{pmatrix} -a \sin u \\ b \cos u \\ \pm c \end{pmatrix} \quad (2)$$

(with parameters $u \in [0, 2\pi[$ and $v \in \mathbb{R}$) are entirely contained in \mathcal{S} .

Solving Eq. (1) for z and parametrizing \mathcal{S} by $(x, y, \pm z(x, y))^T$, the Gaussian curvature of \mathcal{S} can be expressed in terms of the underlying Cartesian coordinate system as

$$K = -\frac{1}{a^2 b^2 c^2} \cdot \frac{1}{\left(\frac{x^2}{a^4} + \frac{y^2}{b^4} + \frac{z^2}{c^4}\right)^2}. \quad (3)$$

The support function d of \mathcal{S} , i.e., the distance of \mathcal{S} 's tangent planes to the origin of the coordinate system is given by

$$\frac{1}{d} = \sqrt{\frac{x^2}{a^4} + \frac{y^2}{b^4} + \frac{z^2}{c^4}} \quad (4)$$

provided that the point of contact is the point $(x, y, z)^T$ whose coordinates satisfy Eq. (1). Thus, we rewrite the formula for the Gaussian curvature given in Eq. (3) in terms of the support function of the one-sheeted hyperboloid and find

$$K = -\frac{d^4}{a^2 b^2 c^2}. \quad (5)$$

This is the analogue to a formula given by Wunderlich in [6] for the Gaussian curvature of an ellipsoid. Actually, this simple formula relating the distance of the tangent planes and the Gaussian curvature is derived for quadrics with center (including ellipsoids, one-, and two-sheeted

hyperboloids) in [4]. However, only for the one-sheeted hyperboloids there is a negative sign.

The support function of \mathcal{S} equals the distance of the surfaces points of \mathcal{S} to the origin exactly at the vertices $(\pm a, 0, 0)^T$ and $(0, \pm b, 0)^T$. Inserting $d = \pm a$ and $d = \pm b$ in Eq. (5) we find $K_a = -a^2 b^{-2} c^{-2}$ and $K_b = -b^2 a^{-2} c^{-2}$. Since $a < b$ by assumption we can easily recognize that the minimum of the Gaussian curvature on \mathcal{S} equals $K_{\min} = K_b$ which is the case at the two vertices $(0, \pm b, 0)^T$.

The iso-curve with $K = K_a$ splits into a pair of congruent ellipses concentric with the quadric \mathcal{S} (see Figure 2) with carrier planes through the x -axis. The length of its semi-minor axis equals a (i.e., the semi-minor axis of \mathcal{S} in the x -axis). The length of the semi-major axis of these two ellipses equals $\frac{1}{a} \sqrt{a^2 b^2 + b^2 c^2 - c^2 a^2}$.

Since a, b, c are constant and with Eq. (5) in mind, we can state:

Theorem 2.1. *The tangent planes of the hyperboloid \mathcal{S} along a curve of constant Gaussian curvature K_0 are at fixed distance $d_0 = \sqrt{abc^4 / -K_0}$ from the origin, and thus, they envelope a sphere which is concentric with \mathcal{S} and has radius d_0 .*

Independent from the choice of a, b , and c , K is always negative, as it was to be expected for a ruled surface without singular surface points. From Eq. (3) we can deduce:

Theorem 2.2. *The curves of constant Gaussian curvature on a one-sheeted hyperboloid with Eq. (1) are the quartic curves of intersection of the hyperboloid with concentric and coaxial ellipsoids \mathcal{E} with equation*

$$\mathcal{E} : \frac{x^2}{a^4} + \frac{y^2}{b^4} + \frac{z^2}{c^4} = \frac{1}{abc\sqrt{-K}} = \frac{1}{d^2}. \quad (6)$$

Figure 2 shows some curves of constant Gaussian curvature on a one-sheeted hyperboloid.

Since $K < 0$ for all points on \mathcal{S} , Eqs. (6) are the equations of ellipsoids with real points. We can reformulate this result as:

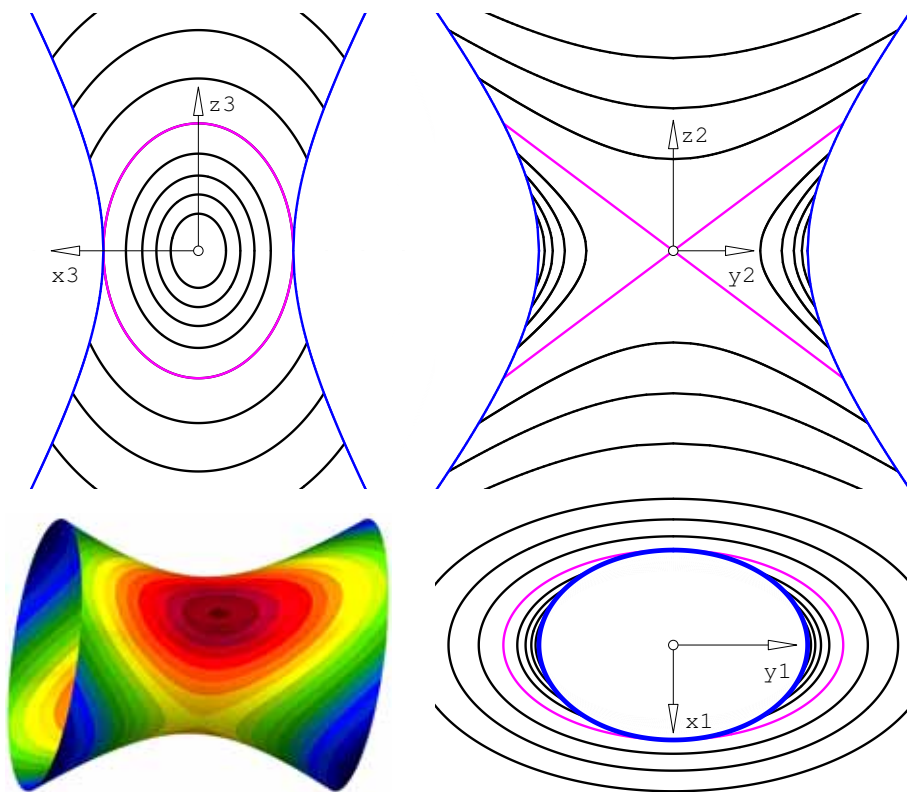


Figure 2: Curves of constant Gauss curvature: right side view and front view (top row, left and right), top view (below, right). The magenta ellipses are the iso-curves $K = K_a$.

Theorem 2.3. *The curves of constant Gaussian curvature $K_0 < 0$ on a one-sheeted hyperboloid are the curves of contact of a developable ruled surface tangent to the hyperboloid \mathcal{S} with Eq. (1) and a concentric sphere of radius $d_0 = 1/\sqrt{abc\sqrt{-K_0}}$.*

This result is similar to that given by Wunderlich in [6] for the curves of constant Gaussian curvature on an ellipsoid.

Figure 3 illustrates the contents of Theorems 2.1 and 2.3. The part of the developable surface joining the curve of constant Gaussian curvature on \mathcal{S} and a spherical curve (on the sphere \mathcal{E} mentioning in Theorem 2.1) is shown.

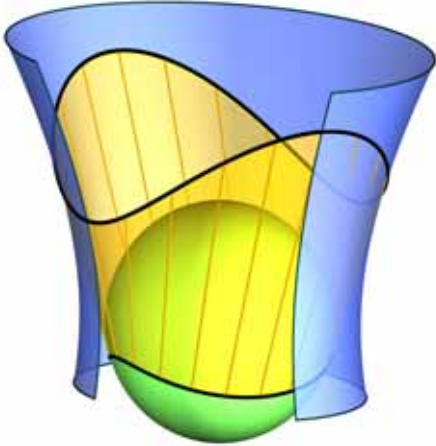


Figure 3: The tangent planes of \mathcal{S} along a curve a constant Gaussian curvature are tangent to a concentric sphere and form a developable surface which is in line contact with both quadrics.

Assume \mathcal{S} is parametrized as one of the ruled surfaces given by Eq. (2) with directrix $l(u) = (a \cos u, b \sin u, 0)$ and the unit vector field $g(u)$ parallel to either $\bar{g} = (-a \sin u, b \cos u \pm c)$. (The $+$ and $-$ correspond to the *right* and *left* regulus.) According to LARMARLE's formula (see for example [2, 4]), the Gaussian curvature can

be computed via

$$K = -\frac{\delta^2}{(\delta^2 + v^2)^2} \quad (7)$$

where $\delta = \delta(u) = \det(\dot{g}, g, \dot{l})(\dot{g}, \dot{g})^{-1}$ is the distribution parameter of the ruling $r(u) = d(u) + v \cdot g(u)$ and the parameter v equals the distance of the surface point (u, v) from the striction point. (The $\dot{}$ indicates differentiation with respect to u .) Obviously, the Gaussian curvature considered as a function on a ruling (or equivalently $K(u, v)$ restricted to $u = u_0$ (fixed)) attains its minimum exactly at $v = 0$ which corresponds to the striction point. Therefore, we can say:

Theorem 2.4. *The curves of constant Gauss curvature touch the rulings exactly at the striction points.*

Figure 4 shows that the rulings and the iso-curves of K are in contact at the striction points.

For a fixed value $K_0 < 0$ the hyperboloid \mathcal{S} (1) and the ellipsoid \mathcal{E} from (6) span a pencil of quadrics passing through the curve with constant Gaussian curvature K_0 . Within this pencil we find four singular quadrics. The first of which is a quadratic cone emanating from $(0, 0, 0)^T$. The remaining three are

$$\begin{aligned} \mathcal{T} : \quad & b^4 \beta x^2 + a^4 \alpha y^2 = a^4 b^4 (c^2 \lambda + 1), \\ \mathcal{F} : \quad & c^4 \gamma y^2 + b^4 \beta z^2 = b^4 c^4 (a^2 \lambda - 1), \\ \mathcal{R} : \quad & -c^4 \gamma x^2 + a^4 \alpha z^2 = a^4 c^4 (b^2 \lambda - 1), \end{aligned} \quad (8)$$

with $\alpha := b^2 + c^2$, $\beta := c^2 + a^2$, $\gamma := a^2 - b^2$, and $\lambda^{-1} := abc\sqrt{-K}$. The equations of \mathcal{T} , \mathcal{F} , and \mathcal{R} as given in (8) are the equations of the top, front, and right side view since they are relations in two variables only. From that we learn:

Theorem 2.5. *The curves of constant Gauss curvature on a one-sheeted hyperboloid are non-rational quartic curves.*

The top view and the right side view of the curves of constant Gauss curvature on a one-sheeted hyperboloid are ellipses. The front view of the curves of constant Gauss curvature on a one-sheeted hyperboloid are hyperbolae.

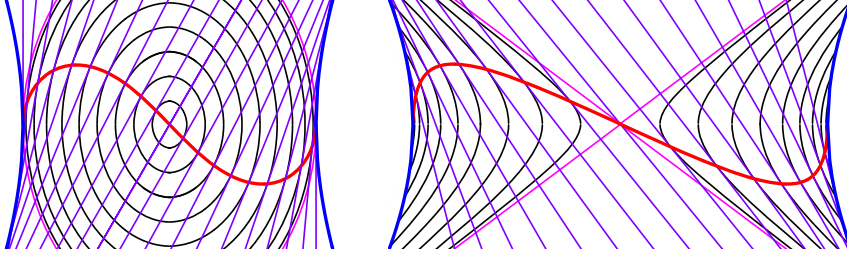


Figure 4: The curves of constant Gaussian curvature (black) touch the generators (violet) exactly at the central points. The central curve (striction curve) is shown in red (right-side view and front view).

Figure 2 shows the three principal views of the iso-curves of the Gaussian curvature. In Figure 4 the right side view and the front view are shown.

3. MEAN CURVATURE

The Mean curvature can be given in two equivalent ways: First, we can start with the parametrization $(x, y, z(x, y))^T$ with z being a solution of Eq. (1). Thus, we have

$$M = \frac{d^3}{2a^2b^2c^2}L \quad (9)$$

with L being a quadratic function in x, y, z :

$$L = (b^2 - c^2)\frac{x^2}{a^2} + (a^2 - c^2)\frac{y^2}{b^2} - (a^2 + b^2)\frac{z^2}{c^2}. \quad (10)$$

From Eq. (9) we can immediately see: The curves on \mathcal{S} with vanishing Mean curvature $M = 0$ are described by Eq. (1) and (10). The latter equation is that of a quadratic cone \mathcal{L} emanating from \mathcal{S} 's center. Eliminating x , or y , or z , we obtain the equations of the top view \mathcal{T} , the front view \mathcal{F} , and the right side view \mathcal{R} of the curve of vanishing Mean curvature. Thus, we have

$$\begin{aligned} \mathcal{T} : b^2\beta x^2 + a^2\alpha y^2 &= a^2b^2(a^2 + b^2), \\ \mathcal{F} : -c^2\gamma y^2 + b^2\beta z^2 &= b^2c^2(c^2 - b^2), \\ \mathcal{R} : -c^2\gamma x^2 - a^2\alpha z^2 &= a^2c^2(c^2 - a^2). \end{aligned} \quad (11)$$

Now we can show the following result:

Theorem 3.1. *The curve m of vanishing Mean curvature on the one-sheeted hyperboloid \mathcal{S} with*

Eq. (1) splits into the pair of smallest circles on \mathcal{S} if, and only if, $c = a$.

Proof. By assumption $a < b$. If $c = a$, the right side view \mathcal{R} given in Eq. (11) splits into a pair of line segments on the lines $x\sqrt{b^2 - a^2} \pm z\sqrt{a^2 + b^2} = 0$. These lines are the views of the projecting planes ζ and $\bar{\zeta}$ with the same equation and they are real. These two planes meet \mathcal{S} in a pair of real ellipses. The semi-major and semi-minor axis of the ellipse m' showing up in the top view \mathcal{T} are of length $a' = \sqrt{\frac{a^2 + b^2}{2}}$ and $b' = b$. The right side view m''' is also an ellipse and its semi-major and semi-minor axis are of length $b''' = b$ and $a''' = \sqrt{\frac{b^2 - a^2}{2}}$. Since $(a')^2 + (a''')^2 = b$, the second principal axis of m is of length b . This shows that m is a circle. All circles on \mathcal{S} are contained in planes parallel to ζ and $\bar{\zeta}$. Since ζ and $\bar{\zeta}$ pass through the \mathcal{S} 's center and the vertices $(0, \pm b, 0)^T$, these two circles are the smallest on \mathcal{S} .

In case of $c = b$ the front view \mathcal{F} given in Eq. (11) equals $y^2(b^2 - a^2) + z^2(a^2 + b^2) = 0$ which is a pair of complex conjugate lines (or planes) since $a < b$. Thus the curve m is the intersection of a pair of complex conjugate planes with \mathcal{S} and carries no real point. \square

In a completely different way Krames has shown in [3] the following: If one point on the smallest circles of a one-sheeted hyperboloid is a point of vanishing Mean curvature, then any

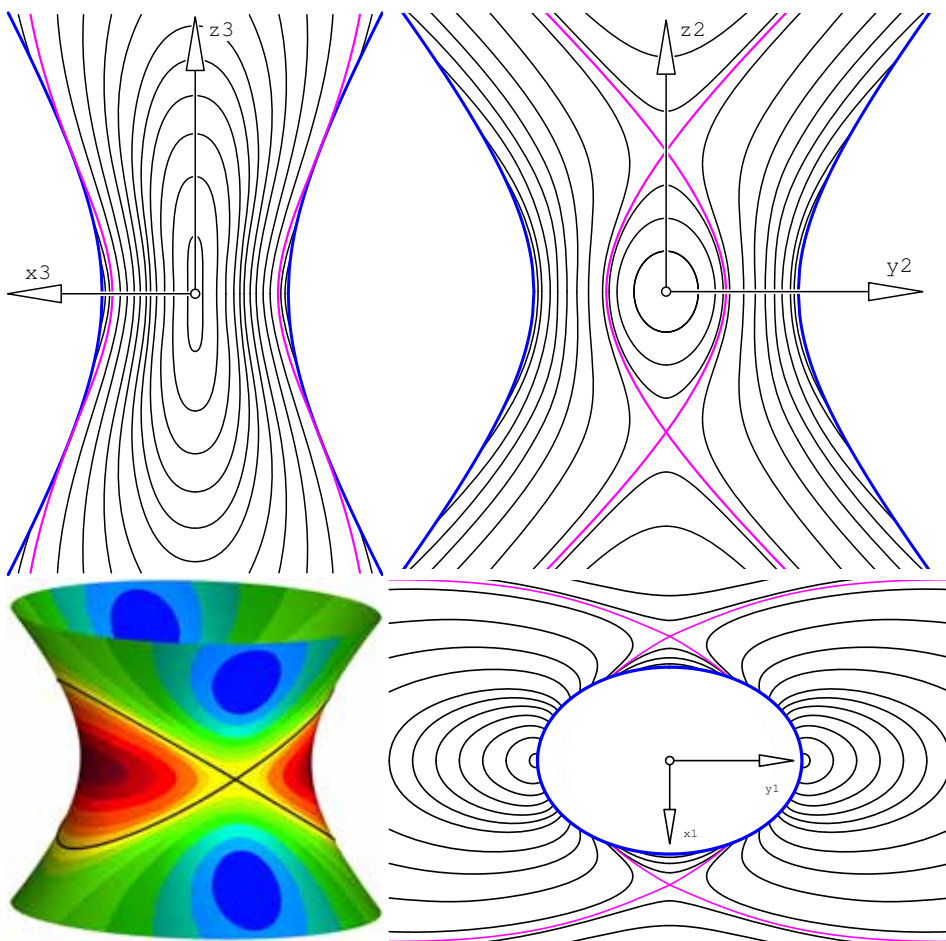


Figure 5: Curves of constant Mean curvature: right side view, top view, front view (top row and bottom row, right) showing the singular curve (magenta) on the hyperboloid $a = 1, b = \sqrt{2}, c = \sqrt{3}$; areas with positive (red) and negative (blue) Mean curvature separated by the (black) ellipses all of whose points show $M = 0$ on the hyperboloid $a = c = 1, b = \sqrt{2}$ (bottom, left).

point on the smallest circles is a point of vanishing Mean curvature.

Figure 5 (bottom row, right) shows a one-sheeted hyperboloid (with $a = c = 1$, $b = \sqrt{2}$) with its curve of vanishing Mean curvature consisting of a pair of circles.

In order to describe the curves of constant Mean curvature we can turn to an expression equivalent to Eq. (9)

$$4a^4b^4c^4M^2 = L^2d^6. \quad (12)$$

From Eq. (12) we can deduce:

Theorem 3.2. *The curves of constant Mean curvature on a one-sheeted hyperboloid are algebraic curves of degree 12.*

The principal views of the curves of constant Mean curvature on a one-sheeted hyperboloid are algebraic curves of degree 6.

Proof. The curves of constant Mean curvature on \mathcal{S} are described by the quadratic equation (1) of \mathcal{S} and the equation of a sextic surface with equation (12). According to Bezout's theorem the curves' degree equals $2 \cdot 6 = 12$.

Each of the principal views is traced twice because of \mathcal{S} 's symmetry. Hence, the degree of the curves appearing in the three principal views reduces to $12 : 2 = 6$. \square

Among the curves of constant Mean curvature there are singular curves if the Mean curvature is either $\frac{bc}{3a^3\sqrt{3}}$ or $\frac{ac}{3b^3\sqrt{3}}$. These singular curves have four additional real double points at

$$\begin{aligned} & \left(0, \pm \frac{b}{\sqrt{\alpha}} \sqrt{3a^2 + b^2}, \pm \frac{c}{\sqrt{\alpha}} \sqrt{3a^2 - c^2}\right)^T \\ & \text{or} \\ & \left(\pm \frac{a}{\sqrt{\beta}} \sqrt{a^2 + 3b^2}, 0, \pm \frac{c}{\sqrt{\beta}} \sqrt{3b^2 - c^2}\right)^T \end{aligned}$$

depending on whether $3a^2 - c^2 > 0$ or $3b^2 - c^2 > 0$. These curves are shown in Figure 5 (magenta curves).

4. PRINCIPAL CURVATURES

The principal curvatures κ_1 and κ_2 are the eigenvalues of the Weingarten mapping ω , cf. [1, 5]. If the Weingarten mapping on \mathcal{S} is described by the quadratic matrix W , then the Gaussian curvature equals $K = \det W = \kappa_1 \kappa_2$ and the Mean curvature equals $M = \frac{1}{2} \text{tr}(W) = \frac{1}{2}(\kappa_1 + \kappa_2)$. From these two equations we can eliminate either κ_1 or κ_2 and find

$$\kappa^2 - 2M\kappa + K = 0 \quad (13)$$

where we suppress the unnecessary indices. With Eqs. (5), (9), and (10) we can rewrite Eq. (13) as

$$a^2b^2c^2\kappa^2 - d^3L\kappa - d^4 = 0. \quad (14)$$

Separating the roots appearing in Eq. (14) and squaring it once again, we obtain an implicit equation, i.e., a polynomial equation, of a surface that intersects \mathcal{S} with equation (1) along the curves of constant principal curvature. We find the implicit equation of an algebraic surface of degree with two disconnected components:

$$(a^2b^2c^2\kappa^2 - d^4)^2 - d^6\kappa^2L^2 = 0.$$

Thus, we have:

Theorem 4.1. *The curves of constant principal curvature on a one-sheeted hyperboloid are algebraic curves of degree 16. The principal views of these curves are algebraic curves of degree eight.*

Remark: The reduction of the degree of the image curves is caused by the symmetry of the surface \mathcal{S} with respect to the image planes of the three orthogonal projections.

Figure 6 shows the right side view and the front view of the two families of curves of constant principal curvature on the one-sheeted hyperboloid with $a = 1$, $b = \sqrt{2}$, and $c = \sqrt{3}$.

The implicit equation of the curves of constant principal curvature may not be useful for drawing or plotting. Thus, we derive a parametrization of these curves. Actually, this parametrization is algebraic (but free of elliptic functions). The

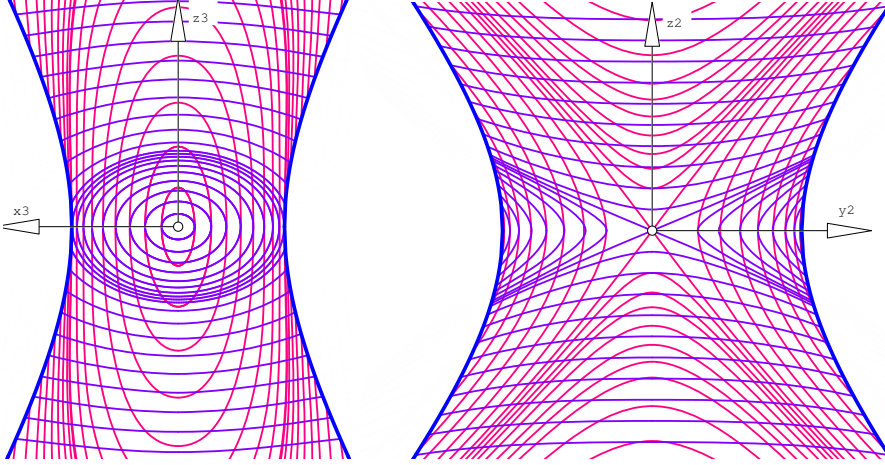


Figure 6: Curves of constant principal curvature (right side view and front view).

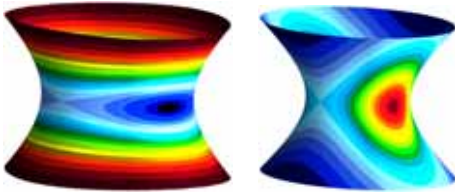


Figure 7: Distribution of principal curvatures on a one-sheeted hyperboloid.

iso-curves of the principal curvature are curves on \mathcal{S} (with equation (1)). We parametrize these curves by d , *i.e.*, the support function of the hyperboloid \mathcal{S} . Thus, d and the coordinates x , y , z of a point on such a curve are also subject to Eq. (4). Further, the coordinates x , y , and z of a point on an iso-curve of κ fulfill Eq. (10). This system of three quadratic equations is linear in the squares of x , y , and z . The solution of this

system of linear equations reads

$$\begin{aligned} x^2 &= \frac{a^4}{\beta\gamma\kappa d^3}(d^3 - b^2 c^2 \kappa)(a^2 \kappa + d), \\ y^2 &= -\frac{b^4}{\alpha\gamma\kappa d^3}(d^3 - a^2 c^2 \kappa)(b^2 \kappa + d), \\ z^2 &= \frac{c^4}{\alpha\beta\kappa d^3}(d^3 + a^2 b^2 \kappa)(d - c^2 \kappa). \end{aligned} \quad (15)$$

Here, we have used the fact that Eq. (14) is linear in L , and thus, we have

$$L = \frac{\kappa^2 a^2 b^2 c^2 - d^4}{\kappa d^3}. \quad (16)$$

Note that the curves of constant principal curvature are different from the (principal) curvature lines. The latter are characterized by the fact that their tangents are always principal tangents. Furthermore, the two one-parameter families of curvature lines are quartic curves and appear as the intersection of the given quadric with the two one-parameter families of its confocal quadrics as illustrated in Figure 8.

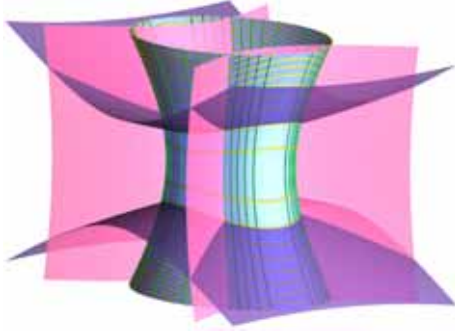


Figure 8: The principal curvature lines do not agree with the curves of constant principal curvature. Principal curvature lines from either family as the intersections of the hyperboloid (light blue) with two confocal quadrics.

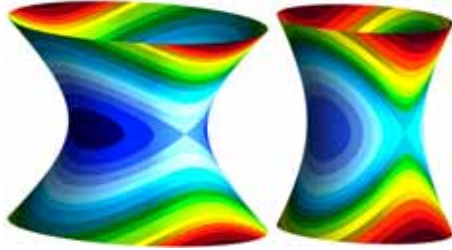


Figure 9: Curves of constant ratio of the principal curvatures on two hyperboloids: $a = 1, b = \sqrt{2}, c = \sqrt{3}$ (left), $a = c = 1, b = \sqrt{2}$ (right).

5. RATIO OF PRINCIPAL CURVATURES

The ratio $R = \frac{\kappa_1}{\kappa_2}$ helps to classify the Dupin indicatrix. If $R = 1$ at some point P on a surface (which will not happen on the one-sheeted hyperboloid), then P is an umbilic and the indicatrix at P is a circle. Since $K = \kappa_1 \kappa_2 < 0$ at all points on \mathcal{S} , we can only expect $R = -1$. In this case, we have $\kappa_1 = -\kappa_2$, and thus, $M = 0$. At such a point the surface behaves (locally) like minimal surface and the indicatrix consist of a pair of conjugate equilateral hyperbolae.

Figure 9 shows the distribution of the ratio

$R = \kappa_1 \kappa_2^{-1}$ on two different one-sheeted hyperboloids.

We aim at a precise analytic description of the curves of constant ratio of principal curvature. For that, we solve Eq. (14) for κ and find

$$\kappa_{1,2} = \frac{d^2}{2a^2b^2c^2} (dL \mp W)$$

with $W := \sqrt{d^2L^2 + 4a^2b^2c^2}$. Now we let $R = \kappa_1 \kappa_2^{-1}$ and derive an implicit equation for the iso-surfaces of R by solving

$$R = (dL - W) : (dL + W)$$

for W and squaring once. Finally, this yields an implicit equation of degree 4 in x, y , and z :

$$a^2b^2c^2(1+R)^2 + Rd^2L^2 = 0. \quad (17)$$

Now we have:

Theorem 5.1. *The curves of constant ratio of the principal curvatures on a one-sheeted hyperboloid are algebraic curves of degree 8. The principal views of the curves of constant ratio of principal curvatures are algebraic curves of degree 4 due to the symmetry of \mathcal{S} with respect to the principal planes.*

Figure 10 shows the right side view, the front view, and the top view of the iso-curves of the ratio R .

6. CONCLUSION

We have computed the iso-curves of several curvature functions on a one-sheeted hyperboloid. The case of a one-sheeted hyperboloid of revolution is trivial, for the iso-curves of all the functions discussed here are parallel circles.

The iso-curves of the Gauss curvature, the Mean curvature, the two principal curvatures, and the ratio of principal curvatures are algebraic curves on the hyperboloid, indeed on any algebraic surface. This is also the case for the three principal views (orthogonal projections onto a triplet of three mutually orthogonal planes, *i.e.*, in this case the three planes of symmetry). The

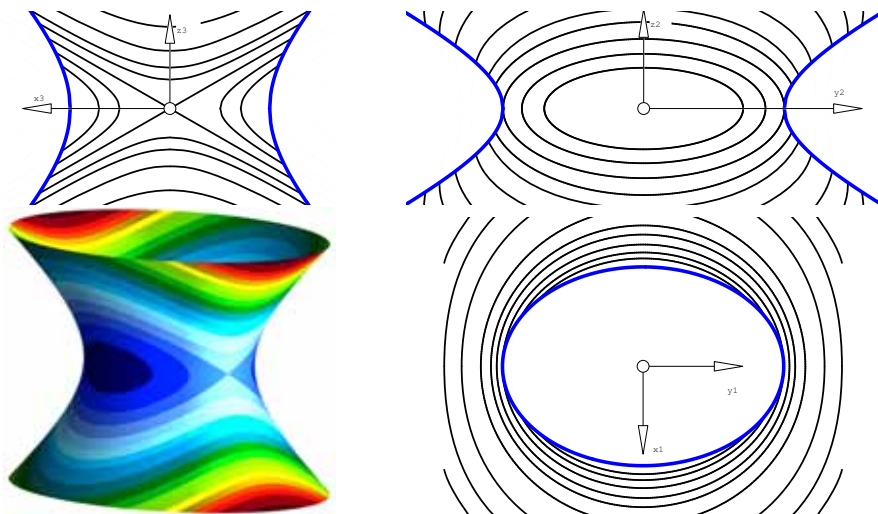


Figure 10: Iso-curves of $\kappa_1 : \kappa_2$ on a hyperboloid ($a = c = 1$, $b = \sqrt{2}$): right side and front view (top row), top view (bottom row, left).

degree of the curves showing up in the principal views are half the degrees of the space curves, since each fibre of any principal projection meets the curve twice.

REFERENCES

- [1] M.P. DO CARMO: *Differential Geometry of Curves and Surfaces*. Prentice Hall, 1976.
- [2] J. HOSCHEK: *Liniengeometrie*. Bibliographisches Institut, Zürich, 1971.
- [3] J. KRAMES: *Zur mittleren Krümmung einschaliger Hyperboloide*. Anz. Akad. Wiss. Wien, math.-naturw. Klasse Nr. 1, 1971, 1–3.
- [4] E. MÜLLER: *Vorlesungen über Darstellende Geometrie*. Band III: Konstruktive Behandlung der Regelflächen. Deuticke, Leipzig und Wien, 1931.
- [5] M. SPIVAK: *A comprehensive introduction to differential geometry*. Publish or Perish, 1978.

- [6] W. WUNDERLICH: *Überblick über die Krümmungsverhältnisse auf dem Ellipsoid*. Festschrift Emil Doležal, 1967.

ABOUT THE AUTHOR

Boris Odehnal studied Mathematics and Descriptive Geometry at the Vienna University of Technology where he also received his PhD and his habilitation in Geometry. After a one-year period as a full interim professor for Geometry at the TU Dresden he changed to the University of Applied Arts Vienna. He can be reached via email at boris.odehnal@uni-ak.ac.at or at his postal address

Abteilung für Geometrie, Universität für Angewandte Kunst Wien, Oskar-Kokoschka-Platz 2, 1010 Wien, Austria.

THE CURVED HORIZON BY LEONARDO: TOWARDS A NEW PERSPECTIVE.

Daniele CALISI

Roma Tre University, Rome

ABSTRACT: For centuries, since it was for the first time clearly represented by Filippo Brunelleschi, linear perspective was addressed and studied by a large number of treatises. The coding took place only after centuries, and the path of knowledge of this science is long and intriguing. But what if one realize that some of the fundamentals of this science are actually misrepresented? In the end, the path that is disentangled over the centuries has taken on board some analysis and put aside other theorizing that perhaps would rather revolutionized the way of drawing in perspective. This is the case under discussion in this short article, to be followed by more detailed studies. Leonardo Da Vinci, in the eighth chapter of the *Trattato di Pittura*, had advanced his doubts in regard to the linear perspective, and its concerns in geometric rules that were rooted to the optical theories on vision. Leonardo was the first to define important precepts, unfortunately not taken into account in the process of codification of perspective science, which introduced a horizon marine, physical and curved, more real and in contrast with the straight horizon of the linear perspective.

Keywords: Geometry, Linear perspective, Leonardo Da Vinci, Curved horizon, Conic curves.

1. INTRODUCTION

In the early years of the Renaissance became more and more the gap between artists and architects in the use of design and techniques related to it. The first ones used the drawing as an artistic tool and painting techniques were taught in art workshops, while the latter ones used primarily as a technical design, according to the technical construction and rendering that were already suggested by Vitruvius.

The painters, driven by an ever more need to faithfully represent reality, embarked on experiments and studies, pursuing the idea of building images that match closer and closer to the real vision.

However, was Filippo Brunelleschi, an architect, that first represented, in two boards, a perspective view of reality. They were lost and constitute the first contribution to the renaissance linear perspective built with the scientific method, which was followed by the abbreviated method by Leon Battista Alberti.

But it must be remembered that the architects of that period continued to represent the architecture without using the perspective, not realizing that this could be an extraordinary method of project management.

Linear perspective is rather widespread in the workshops of the painters, that, pursuing the pictorial realism, applied perspective to the representation of lighting, shadows and chiaroscuro.

The vision of the Renaissance was totally anthropocentric and placed the human eye as the center of projection, determining perspective images by a process of projection and section: a single center of projection, then, whose projective rays hit objects in the scene.

The encoding process of the prospective method is very long and spanning several centuries, including small steps and turns due to glaring oversights and errors of concept: without going too much into the process, repeatedly examined by many, remember that the actual encoding is made to coincide with

the discussion of Gaspard Monge in his texts, at the beginning of the nineteenth century, hence the name of Monge projections for methods of representation.

However, this nominal extension is somewhat anachronistic, since it does not take count many small factors and contributions that have, perhaps more than the disquisitions of Monge, helped to know deeply the science of perspective, and nor includes many other contributions that have dealt with the perspective under other interesting aspects, often gone unnoticed for centuries.

Definitely Leonardo da Vinci is one of the authors who gave, in the course of his life, more hints for thought over, analysis, specific studies and practical experiments.

2. THE AVANT-GARDE OF LEONARDO DA VINCI

Leonardo filled his notebooks of information related to all sorts of thoughts, sketches and reasoning about different issues or information of daily life.

A collection of these manuscripts made between 1482 and 1518, recovered and re-arranged by Francesco Melzi, became the *Trattato di Pittura* (Treatise on Painting), which appeared for the first time in Paris in 1550 and in 1651 in a very short form compared to the original, while an other Italian edition appeared in Naples in 1733.

The *Trattato di Pittura* is inserted in Cod. Vaticano Urbinate 1270 and is divided into 8 sections, each of which consists of a set of prepositions.

The treaty was later re-released also in 1815, transcribed and commented by Guglielmo Manzì, complete with twenty-two tables with 221 illustrations.

Leonardo does not encode the perspective technique, but as for each of his reasoning he gives some precepts, explains its practical experiments, analyzes, nevertheless never advancing in mathematical calculations or explanations of geometric theorems that were

probably clear and obvious for him. Nevertheless, it needs to read behind the lines of its prepositions even with the help of the useful diagrams beside the text.

In particular, among the many prepositions provided in the Treaty (all very useful in hindsight), there are some in which Leonardo enters a difficult topic, unique and innovative for its time in which the few known perspective rules were those by Alberti: the curved horizon. In the second half of the '400 it were trying to find the rules for the exact construction of the perspective in order that it might be the most likely one compared with the ocular vision; However, they pursued this result by applying some abstractions derived from millennial arguments on the optics, and more specifically on the straightness of visual rays.

While the basics of the science of perspective were founded on this "soft ground" because it was based on assumptions not entirely true (although the visual rays are straight, not the same can be said of the course of lines on the terrestrial calotte), Leonardo lingers instead to think about the true horizon, what we perceive looking at the sea, which is not exactly a straight horizontal line.

In such a context, the topic of Leonardo was definitely not foregone, nor trivial.

Yet one can not say that Leonardo did not know to draw in perspective according to the canonical rules, as evidenced by all his works (not to mention that wonderful pictorial artifice that is the aerial perspective), and in particular a simple preparatory drawing for the Adoration of the Magi (1471 ca., Gabinetto Disegni e Stampe degli Uffizi, Florence, 436E) makes us clearly understand what skill Leonardo da Vinci had in the perspective art.

However, the spirit of Leonardo was as a true "scientist", a careful observer of reality, to which nothing escaped, from the movement of currents and fluid to the course of the winds, from the harmonies of the geometries to the vivisection on dead bodies. His greatest art was to be able to "look", that is to relate to the

world around always with curiosity and analytically way.

Leonardo evidently could not refrain from a deeper reality and a constantly perception, in front of his eyes, of a curved horizon that denied the same way he draw (more didactic than slavishly followed the rules of perspective).



Figure 1: L. Da Vinci, studing sketch for Adorazione dei Magi.

3. THE HORIZON IN ACCORDANCE WITH LEONARDO

His reflections about flow in the eighth section of the Trattato di Pittura which is titled "*Del Orizzonte*" and consists of 7 facades ranging from f. 283 r to f. 286 r.

It is a very short text in which are collected the last considerations of Leonardo on the horizon. The text is composed of more parts indicated by the letters *a* to *K*, where Leonardo discusses the innovative and thorny theme of curved horizon, intended as the physical horizon to the rules of perspective; it is divided into several titles: "*Dell' orizzonte*" (*f*), "*Del vero orizzonte*" (*g*), "*Dell' orizzonte*" (*H*, ma segnato in seguito sul quarto rigo del paragrafo), "*De l'orizzonte*" (*i*), "*Se l'Occhio che vede l'orizzonte maritimo stando coli piedi alla pelle d'esso mare vedo esso orizzonte piu basso di se*" (*K*), "*Del'orizzonte specchiato nell'acqua corente*", "*Dove l'Orizzonte e' specchiato nell'onda*" e "*Perche l'aria grossa vicina all'orizzonte si fa rossa*" (the last three have no reference letters).

These prepositions were not followed by a

coding, as happened for the *perspectiva artificialis*, even if Leonardo laid the foundation for a real coding of the representation of perspectival space that actually followed geometrical and not empirical rules.

In the first preposition Leonardo already shows a deep knowledge of the geography and geometry of the Earth, making it clear right away that there are multiple horizons at different distances from the eye, because their distance is closely related to the height from which these horizons are observed, if the sea water level, whether from human height, or if top of a mountain, along the perpendicular to the ground (towards the center of the earth). However, Leonardo says, "*...quelli che sono infra terra non hanno l'orizzonte con eguale distanza, perchè la superficie della terra non è egualmente distante dal centro del mondo, onde non è di perfetta sfericità, com'è la pelle dell'acqua; e quest'è causa di tal varietà di distanze infra l'occhio e l'orizzonte...*", introducing the theme of the irregular course of the Earth's surface, defined, centuries later, as a geoid.

Continuing in the first preposition Leonardo tries to prove, starting from a more discursive text to true geometric proofs, that the visual pyramid embraces an infinite space and demonstrate it also through the Alberti rules of projection of visual rays and section with the picture plane.



Figure 2: L. Da Vinci, First sketch on the chapter 8 of the Trattato di Pittura (f. 283 v). The visual pyramid sectioning the picture plane.

As the eye sight points more and more distant from the observer, it seems that such points

projected onto the picture plane have a higher position, as, in contrast, a bird which fly away from the observer according to a straight line, it seems instead that he assumes positions always lower (analysis of Leonardo in Manuscript K, f. 121 r). Besides his statements echo the tenth and eleventh theorems of Euclid: the more distant parts of a plane that is located below the eye seem higher while the more distant parts of a plane that is located above the eye appear lower.

In addition, Leonardo realizes that the visual rays intercept the Picture plane in points ever lower than the height of the eye and whose mutual distance is always smaller as the eye sights more distant points, that is this distance may tend to infinitesimal values. The intersection may be the same height only in the moment in which the visual ray is parallel to the ground, i.e. when space bs is infinite, by sighting a point at infinity.

Leonardo was able, in a few sentences, to quibble with the concept of infinitely great, the infinitely small, and the representation of an infinite point (point improper) on the Picture plane, anticipating by centuries the full maturation of the concept of vanishing point by Desargues.

4. PROSPECTIVE SPACE AND NATURAL SPACE

In the preposition *e* Leonardo introduces a very important concept, which we'll be basic in studying the curved horizon.



Figure 3: L. Da Vinci, second sketch on the chapter 8 of the Trattato di Pittura (Urb. Lat. 1270, f. 284 r).

In perspective view, things that are a little

smaller than the height of the observer are very close, but in the example proposed by Leonardo the object *at* which sight the natural horizon in *b*, will see the object *ru* equal to his same height as *a*, *r* and *b* are aligned in the hypotenuse of the same pyramid (cone). But soon he would point out that this pyramid is not the one that determines the perspective, because in this case it should have infinite space in front of it, when in this case between *a* and *b* there are just 7 miles.

The theme of the perspective pyramid was well known and studied, and even in the classical period Titus Lucretius Carus, in the fourth chapter of the *De Rerum Natura*, cites it to explain the perception of a regular column portico seen from the inside, which gradually contracts in the vertex of a narrow cone, joining the roof to the ground and the whole right side to the left, until unites them in that dark vertex of the cone. Dark probably because it was not yet defined the representation of the vanishing point.

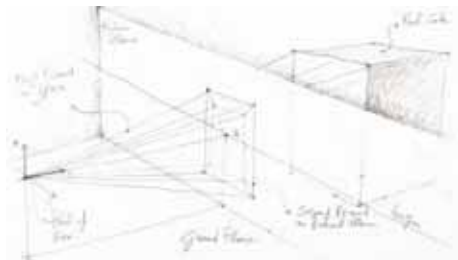


Figure 4: First pyramid has vertex on the point of view and base non defined. Second pyramid is on the picture plane, and is vertex is O_0 , projection of the point of view on the picture plane.

Leonardo himself speaks better about this pyramid in Manuscript G, clearly defines two of them, one which has a visual center in the eye and remote base in an infinite space, and the second which is the perspective ones that has the base in the eye and vertex in the horizon. This second pyramid is the one that is created in the picture plane by determining the perspective view, and is therefore entirely flat.

If it were not flat then we'll be looking at something different, an anamorphic illusion of space, such as the false colonnade of Borromini in the Palazzo Spada.



Figure 5: When the second pyramid is on space, there will be an optical illusion as for Borromini's corridor for Palazzo Spada, in Rome.

Leonardo, however, in the last sentence says that of the perspective pyramid *non si da in pratiche* (we don't have to care about) because it has an infinite space in front of itself, while the natural pyramid has seven miles. This demonstrates the doubt of the scientist towards to linear perspective, resulting in a clearly marked difference between perspective space and natural space and, at the same time, he rooted the limits of a science that did not really represent the physical space.

As a further demonstration of the skepticism of Leonardo, the step *g* in which at the same time he demonstrates the rules of artificial perspective, analyzes and finally refutes them.

After demonstrating full awareness of the vanishing point in step *f*, in the section *g*, "*del vero orizzonte*" (*on the true horizon*), Leonardo introduces, for the first time, the concept of the horizon as a line of contact between two horizontal planes, the earth and the sky: the centric line as it was defined by Alberti, but that no one had yet made to coincide with the sea horizon. This line is situated at the eye height without any doubt, obviously if the earth and sky were two parallel planes and if they were of infinite size. The hypothetical tone of Leonardo is not out of place because he doubts the perspective rules, so much so that in the next step brings us back to reality by showing that the prospective horizon can not be a one-

centric line that passes through the eye, because even if sky and earth were two parallel planes, the earth does not have the same size than the sky.

The step is extremely important because for the first time is made a distinction between the perspective horizon and natural horizon.



Figure 6: L. Da Vinci, fifth sketch on the chapter 8 of the Trattato di Pittura (Urb. Lat. 1270, f. 284 v).

The space of the sky is represented by *ba* projecting its image on the picture plane in *m*. However the more the point moves away the more the horizon coincides with the height of the point of view *g*. The land, instead, has a limited physical size *ef* which projects its image in *n*, which is more or less at the navel of a man.

Since, therefore, heaven and earth can not be parallel planes, but are one convex and the other concave, it is clear that there are infinitely many natural horizons (dependent on the height of the eye) that will never be at the same height of the observer: so the natural is the true horizon. Among other things, since the earth is spherical, the lines themselves do not have much sense, since the only way to join two points of equal height is an arc of circumference: the heights of things will have to be calculated not from a horizontal plane, but from the center of the earth, according to a straight line to the earth's surface in each point.

5. ON THE TRUE HORIZON

In the preposition *K*, "*Se l'occhio che vede l'orizzonte maritimo stando coli piedi alla pelle d'esso mare vedo esso orizzonte piu basso di se*", Leonardo express all his thoughts about the great duality between the perspective horizon and natural horizon, that I best explain later. First, as I just explained, the fact that the

horizon as was usually traced could not be the real one, because it implied parallelism between the surface of the earth and the celestial sphere. The true horizon is actually lower than that of perspective representations and depends on the sphericity of the earth and from the height of the observer. If the geometric horizon is a horizontal line on the picture plane and the fugue of all the horizontal planes which meet at infinity on the line at infinity, the horizon of the sea (as defined by Leonardo) is quantified at a distance of seven miles from the observer, and turns out to be lower on the picture plane.

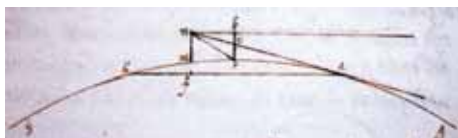


Figure 7: L. Da Vinci, seventh sketch on the chapter 8 of the Trattato di Pittura (Urb. Lat. 1270, f. 285 r): "On the true horizon".

nm is the position of man, br is the picture plane and the curvature of the earth is given by grh , and the natural horizon is a . Since afk is a horizontal segment Leonardo states that the horizon is lower than the foot of nm by an amount equal to mf , and that is lower than the geometric horizon of all bo .

Leonardo in his demonstration, however, is forced to place a horizontal plane (afk) parallel to the celestial vault, and that is a plane that passes under the feet of nm by an amount equal to mf , so that the horizontal plane cross the natural horizon in a .

In this way it can define in the same drawings, both the prospective horizon in b , centric line as defined by Alberti, representation of the intersection at the infinite between the horizontal plane and the sky, and the real horizon in o that is lower than the centric line by an amount equal to bo .

Therefore, while the perspective horizon represent the infinite, and will never be reached "*se non una linea parallela*", the sea horizon is

the limit of the visible surface of the sea, and is only at seven miles if b is at the height of a man.

Leonardo introduces pretty much what today is referred to as "*depression of the horizon*" in celestial navigation, i.e. the angle between the horizontal plane to the eye of the observer O (containing the apparent horizon) and the tangent t in the initial of the curve of refraction for O and which is tangent to the surface of the sea. However, Leonardo clarifies its prepositions always through sketches in side view, and not using at least a second orthogonal view, so we do not know what happens on the picture plane, and how the science of perspective evolves as a result of its reasoning.

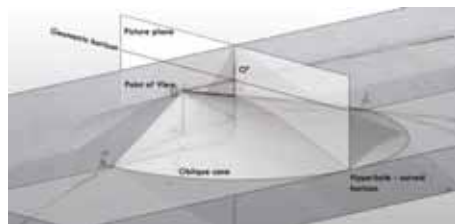


Figure 8: The Leonardo scheme to find the true horizon overlaid to the 3d model.

In my testing, thanks to the use of the computer graphic, I have reproduced the scheme of Leonardo in three dimensions, making it clear and orbitable. In the 3D model it is clear that the marine horizon is actually lower than the geometrical one, but it is not represented by a straight line, as could mislead in the scheme of Leonardo that is not explanatory of these points, which are not actually present in any of his writings. Therefore it can not actually knows if the author was aware or not of the perspective representation of the sea horizon.

What type of curve is the horizon of the sea, and how it's possible to determine it geometrically or parametrically was one of the most important points in this research, still in progress. First of all we analyze the geometric components of the scene: there is a portion of a sphere, symbolizing the earth, whose basic

geometry is a circle; there is a vertical picture plane; a point of view at a certain distance from the picture plane.

Imagining to apply the rules of perspective, what we need to do is conduct a bundle of projecting lines from the point of view to the edge of the circle of the earth, resulting in an oblique cone, the axis of which begins in the center of the circle and with the vertex in the point of view. This cone is sectioned by a vertical plane, i.e. the picture plane.

All elements remind to one of the most important theorems of geometry that allowed the coding and classification of conic curves: the theorem of Apollonius.

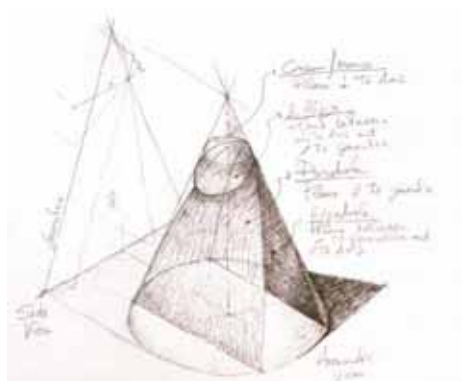


Figure 9: the conic curves according to the theorem of Apollonio

The conics can be determined based on geometrical rules of section of a cone (rectum in the theorem) with a plane that takes some special and limit positions: a plane perpendicular to the axis generates a circumference; a plane with a position between the perpendicular to the axis and a generatrix of the cone generates an ellipse; planes parallel to the generatrix of the cone generate the parabola; finally planes comprised between the generatrix and the axis of the cone itself, generate hyperbolas.

The theorem of Apollonius can in fact be extended to an oblique cone which, in this case,

is sectioned by a vertical plane coinciding with the picture plane: the corresponding conic is precisely a hyperbola located lower than the geometric horizon. Also interesting is to note that the center of the asymptotes of the hyperbola is the same as O_0 , projection on the picture plane of the point of view which also coincides with the vertex of the oblique cone.

The geometry of the parallels and meridians of the globe, ultimately corresponding to the lines running to the vanishing point in the Alberti diagram, is quite similar to what we have just defined: the parallels create oblique cones similar to the one that generates the horizon, thus producing still hyperbolas.

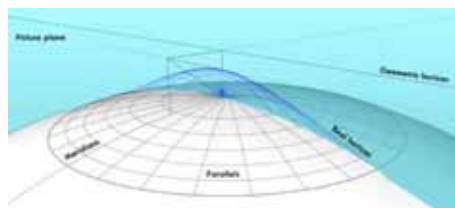


Figure 10: 3D model of the earth as for the Leonardo Scheme. In blue the perspective representation of meridians and parallels.

The geometry of the meridians is affine, but in this case the base of the cone is vertical and coincides with a maximum circumference of the earth, and its vertex is always in the point of view. The plane section is always the picture plane, therefore also the meridians are represented on the panel by hyperbolas.

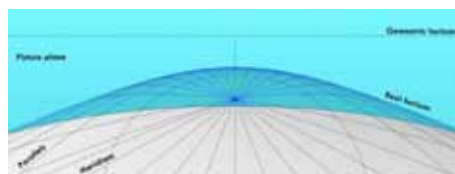


Figure 11: From the point of view, facing the picture plane, and watching the point a . With this conditions, the perspective representation of meridians and parallels is coincident with the real earth behind the picture plane.

6. CONIC CURVE BY APOLLONIO AND THE GEOMETRY OF THE CURVED HORIZON

Once one has determined the basic scheme for the identification of the reference conic, we must impose some basic assumptions useful for demonstrating the generic parametric formula.

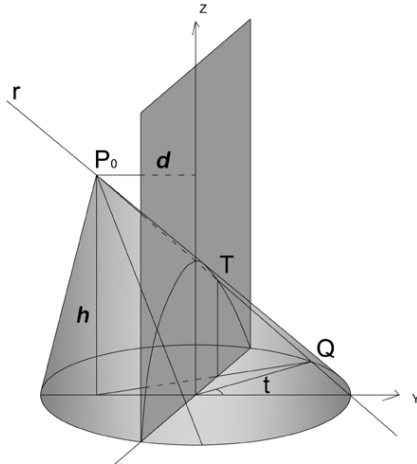


Figure 12: Graphic scheme used to find the parametric equation of the generic hyperbole.

First we determine some of the coordinates of the points to which the straight line r (generatrix of the cone) passes, defined not by numeric values but by generic parameters: P_0 (0, -d, h), T (x, 0, z), Q (cos t, sin t, 0). The point P_0 is located at a distance d from the z vertical axis, and at a distance h from the y horizontal axis. Other two imposed hypotheses provide, as from the scheme, that the distance d is between 0 and 1 ($0 < d < 1$), and that h is greater than 0 ($h > 0$).

The point Q on the circumference of the base is determined by coordinates that depend on the trigonometric values associated to angle t . Instead, the point T of intersection of the straight line with the picture plane has unknown coordinates for the moment, obviously dependent from the other two points that define the equation of the r line,

considering the position of a generic point S belonging to the same straight line. This point has for coordinates: $x(s) = (1-s)\cos t$; $y(s) = -sd + (1-s)\sin t$; $z(s) = sh$.

However, to this coordinate system we can already impose some limitations, as for example the fact that although S is a generic point of the straight line r , in reality we need a particular position, i.e. as S coincides with T , and thus when the $y(s)$ is equal to zero.

$$y(s) = -sd + (1-s)\sin t = 0 \quad (1)$$

$$x(s) = (1-s)\cos t \quad (2)$$

$$z(s) = sh \quad (3)$$

From z we get:

$$s = \frac{z}{h} \quad (4)$$

While for the x starting from (2) and inserting the value of z from the (4), we get:

$$\frac{x}{1-z/h} = \cos t \quad (4)$$

$$\frac{h}{h-z}x = \cos t \quad (5)$$

$$\sin t = \sqrt{1 - \left(\frac{h}{h-z}x\right)^2} \quad (6)$$

From the solutions found we can replace s and $\sin t$ in the first equation (1), and going to solve.

$$-\frac{z}{h}d + \left(1 - \frac{z}{h}\right)\sqrt{1 - \left(\frac{h}{h-z}x\right)^2} = 0 \quad (8)$$

$$\sqrt{1 - \left(\frac{h}{h-z}x\right)^2} = \frac{d}{h} \cdot \frac{h}{h-z}$$

$$\sqrt{1 - \left(\frac{h}{h-z}x\right)^2} = \frac{dz}{h-z}$$

$$\begin{aligned}
1 - \left(\frac{h}{h-z} x \right)^2 &= \frac{d^2 z^2}{(h-z)^2} \\
(h-z)^2 - h^2 x^2 &= d^2 z^2 \\
h^2 x^2 + d^2 z^2 - h^2 - z^2 + 2hz &= 0 \\
h^2 x^2 + (d^2 - 1)z^2 + 2hz - h^2 &= 0 \quad (9)
\end{aligned}$$

Since the generic equation of a conic section on the xz plane is defined by the following expression:

$$ax^2 + 2bxz + cz^2 + 2dx + 2ez + f = 0 \quad (10)$$

Comparing (9) and (10) we can derive the respective values:

$$a = h^2, b = 0, c = d^2 - 1, d = 0, e = h, f = -h^2$$

Actually in order that the general equation of a conic section (10) corresponds to a hyperbola must necessarily occur the condition that b^2 is greater than the product of a and c.

$$b^2 > ac \quad (11)$$

i.e. for our specific case:

$$0 > h^2(d^2 - 1) \quad (12)$$

This condition always occurs, because in our hypothesis we imposed that the d was between 0 and 1, therefore its square will be a number minor than 1. From which $(d^2 - 1)$ is necessarily a negative number, which multiplied by h^2 will result a negative number again. Thus, the value 0 will be inevitably greater than the second negative term.

This shows that the horizon of Leonardo is a conic and more precisely an hyperbole.

In fact we could also further develop the equation, inserting inside the numerical values taken from the same drawings or from some size indications that Leonardo himself introduces in the first proposition: " *...perché*

l'occhio, posto alla pelle del mare quieto, vede esso orizzonte vicino un mezzo miglio o circa; e se l'uomo s'innalza coll'occhio, quant'è la sua universale altezza, l'orizzonte si vede remoto da lui sette miglia, e così in ogni grado di altezza scopre l'orizzonte più remoto da sé, onde accade che quelli che sono nelle cime degli alti monti vicini al mare vedono il cerchio dell'orizzonte molto remoto da loro...".

Leonardo evaluate the distance of the sea horizon in excess that would be, according to him, about seven miles for a man of a height of about 1.74m. If one consider the mile Florentine in that period, equal to 1653.61m, which is shorter than the current geographical or nautical mile, the value that comes out is 11570m.

Leonardo's drawing, however, is only a scheme and actually it don't represent the real earth's surface and distances. Therefore we can scale the Leonardo's drawing or based on the value of the height of the man of 1.74m, or based on the distance to the horizon of 11570m, than finding the corresponding values of h and d , so far left as parameters in the equation of hyperbole.

Another solution is instead to recreate the real conditions and distances, apart from the scheme of Leonardo, and determine the corresponding equation of the hyperbola.

However, the representation of the hyperbola that derives from this last experiment, it appears to be not graphically appreciable because its eccentricity makes it very close to a horizontal line, perceiving the curvature only moving very far away from the point of view set by Leonardo.

Anuway the information very relevant and innovative is to have determined that the curved horizon by Leonardo is an hyperbole, independently from the values that, as Leonardo himself reminds us, are quite changeable, because the earth is not perfectly spherical.

There are also other important points to explain, still in the research stage to be honest, and introducing new and important codes for the

perspective representation. I have repeatedly stated that Leonardo was a connoisseur of the perspective art and its rules, so I am convinced that he had noticed how the draw of a curved horizon was divorced from those rules.

In the canonical perspective the horizon is the geometric locus of the vanishing points of all the horizontal lines, finite representation of an intersection of two parallel horizontal planes: the ground plane and a plane parallel to it through the point of view. Similarly, a vanishing point is the representation of an improper point, and its image is determined simply by passing from the point of view a line parallel to the straight line r , until it intersects the picture plane at a point, defined as the vanishing point of the line r and all the parallels to it.

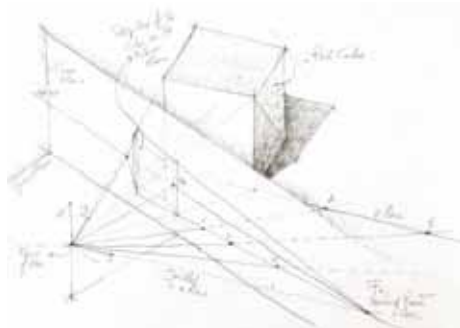


Figure 13: The way to find a vanishing point in the linear perspective.

The perspective of course is based on a process of projection and section: projecting rays connect a point P in space with the point of view, and in the intersection with the picture plane determine a point P' that is its perspective image.

In conclusion we can say that projecting an object from the point of view on the picture plane, we obtain its image, while to obtain fugues of planes and straight lines (which are, respectively, lines and points) we must do pass a plane or a straight line parallel to them, passing through the point of view, and

intersecting with the picture plane.

The horizon of Leonardo do not follow these rules. The first anomaly is that the horizon in the codified perspective is the representation of something that is going to infinity, while the horizon Leonardo is at a finite distance, calculated by the scientist at seven miles from a man tall 174cm.

Just the fact that this horizon can be measured, and especially variable because it depends on the height from which one observes it (if from the sea level or from the top of a mountain), negates the very concept of infinite horizon.

In addition Leonardo determines the curved horizon through a process of projection from the point of view and section with the picture plane, i.e. determines the perspective image of the circumference edge of the Earth. However, the strange thing of all is that the perspective image in this curved perspective also coincides with the geometric horizon.

In fact, the curved horizon of Leonardo can not be understood as a locus of vanishing points, but simply as visual limit to human sight.

This is also proven by the fact that the terrestrial meridians in their projection on the picture plane did not have fugue on the horizon. Their projection is itself a hyperbola with a branch that goes even beyond the visual limit of the curved horizon.

A building of a certain height will curved edges have at the base and at the top (arcs or portions of meridians in special cases) and their perspective representations are two hyperboles that go beyond the visual limit theorized by Leonardo and that meet at the infinite.

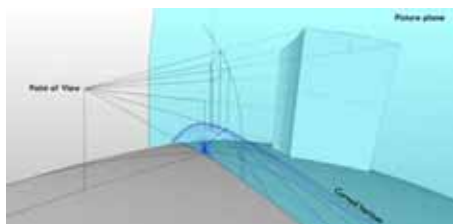


Figure 14: Perspective representation of a building on a spherical surface.

But do not be confused, their improper point of intersection is not traceable on the picture plane at this time, as is feasible for the vanishing points (as explained), because the two hyperbolas are already on the plane, and are the representation of two concentric circles the second of which has a radius equal to the first plus the height of the building: while two parallel lines can meet in a point at infinity, two concentric circles can never meet.

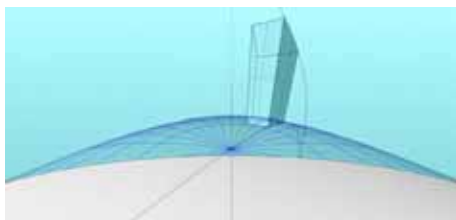


Figure 15: From the point of view facing the picture plane. The sides of the building are represented by hyperbolas, and they never cross.

On the other hand I do not believe that Leonardo was confused on this issue, because, as we have already discussed earlier in this article, is himself to introduce the concept of perspective representation of an improper point in the first preposition of the eighth chapter.

7. CONCLUSIONS

The curved horizon of Leonardo is still one of the most fascinating and least studied theme of descriptive geometry, which, however, shows once again how innovative was his thinking, which I demonstrated the avant-garde more than once in other sectors, from the Lambert law to the radiosity algorithms in modern computer graphics. Leonardo in his writings had given to science many concepts, which if well received would certainly have introduced changes in the visual arts, from painting to geometry to the science of perspective. Unfortunately, his disquisitions went almost completely unnoticed in spite of linear perspective, which was studied, implemented in

the rules, and more and more consciously aware in the minds of theorists, but was completely unaware of the true horizon feared by Leonardo. So that the same John Ruskin, deep connoisseur of the perspective rules and great artist, he says, on page 9, referring to the horizon: *"this line is of great practical use, representing the level of the eye of the observer all through the picture...if there is a horizon to be represented in your picture, as of distant sea or plane, this line defines it"*.

We can not know how it would be the development of science perspective if the precepts of Leonardo had had more followers, however, we can once again praise a great artist and scientist who has always demonstrated a great forerunner of the times, one of the few with the ability of being able to change the history thanks to his enlightening ideas and analysis.



Figure 16: The horizon represented by the lake at the very end of the painting, is at a lower level than the viewer's eye level. L. da Vinci, Gioconda, 1503 - 1515.

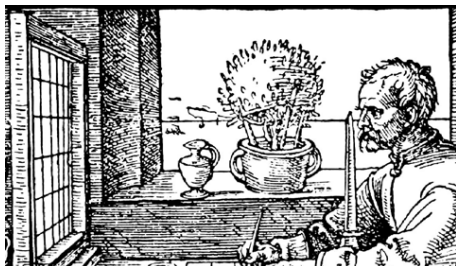


Figure 17: L'occhio del pittore è ad un livello superiore rispetto a quello dell'orizzonte marino, come descritto da Leonardo nel capitolo 8. Xilografia di Albrecht Dürer, in *Underweysung der Messung*, Norimberga 1525, libro 3, fig. 67.

REFERENCES

- [1] L. Da Vinci. *Trattato di Pittura*. Various text, written between 1482 and 1518, matched and ordered by F. Melzi. Later edition.
- [2] *Leonardo da Vinci Libro di Pittura*, edited by Carlo Pedretti and critical transcription by Carlo Vecce. Firenze. Giunti 1995
- [3] V. Valerio. *L'Orizzone e l'Infinito di Leonardo*. In IKNOS, *Analisi grafica e storia della rappresentazione*. 11-50. 2005
- [4] D. Calisi. *Luce ed ombra nella rappresentazione. Rilettura storica e sperimentazioni eidomatiche*. On platform PADIS at the link <http://hdl.handle.net/10805/1047>
- [5] A. Bovi. *Il periodo milanese di Leonardo e la nuova prospettiva di luce e di ombra*. In "Arte Lombarda". Milano. 7/1962, pp. 43 – 48.
- [6] A. Agostani. *La prospettiva e le ombre nelle opere di Leonardo da Vinci*. Domus Galileiana. Pisa 1954. pp. 45.
- [7] J. Ruskin. *The Elements of Perspective, arranged for the use of Schools and intended to be read in connexion with the first three books of Euclid*. Smith, Eider and Co. London 1859.

- [8] D. Calisi. *La modernità delle intuizioni. I colori apparenti nell'interazione tra Superfici*. In Atti della IX Conferenza del Colore, Firenze il 19 e 20 Settembre 2013. Colore e Colorimetria. Contributi Multidisciplinari. Vol. IX A. ISBN 978-88-387-6241-3. Pp 661 - 671.
- [9] D. Calisi. *Elogio della pratica: intuizioni premonitrici ed errori storici*. In Atti del IX Congresso UID Unione Italiana Disegno, Roma 13, 14, 15 Dicembre 2012. Volume "Elogio della teoria. Identità delle discipline del disegno e del rilievo". ISBN 978-88492-2519-8

ABOUT THE AUTHORS

1. Calisi Daniele has a PhD in Science of Surveying and Representation at La Sapienza University of Rome with a thesis that seeks to heal the historical stages and development of the theory of shadows and chiaroscuro and with which he won the National Award 2007, for the best PhD research thesis. He is for years an adjunct professor at the university of Roma Tre where he worked on design, survey (direct and instrumental), digital representation and descriptive geometry, with particular attention to the historical reappraisal of the major theoretical and encoders. Graduated in photography, reportage and photojournalism, it is also sensitive to new technologies for instrumental survey, carrying out research in the field of. At the same time it is also active in several industry research on color and the city, photomodeling, and the representation of the historical city, paying attention in particular to the complex and layered urban fabric, as well as portions of the city lost for different causes. He has published in trade journals and has participated in national and international conferences. He is active in several exhibitions in Italy and in New York on issues of social reportage, using digital techniques.

www.danielecalisi.com, d.calisi@gmail.com

CURVES OF CENTRES OF CONIC PENCILS IN PSEUDO-EUCLIDEAN PLANE

Mirela KATIĆ ŽLEPALO
Polytechnic of Zagreb, Croatia

ABSTRACT: The aim of this paper is to show that in pseudo-Euclidean plane the curve of centres of an order conic pencil is a conic. Also, the conditions on coefficients of base conics are investigated in order to determine which type of conic will be obtained as curve of centres. In addition, it is shown that the curve of centres of a class conic pencil is a line.

Keywords: pseudo-Euclidean plane, conic sections, conic centre, conic pencil

1. BASIC DEFINITIONS

A *pseudo-Euclidean plane* is a real projective plane where the metric is induced by a real line f and two real points F_1 and F_2 incidental with it, [5].

Conics in pseudo-Euclidean plane are divided into ([4]):

- hyperbola intersecting the absolute line in two real and distinct points
- ellipse intersecting the absolute line in a pair of conjugate-imaginary points
- parabola touching the absolute line
- special hyperbola intersecting the absolute line in two real and distinct points out of which one is an absolute point
- special parabola touching the absolute line in an absolute point
- circle intersecting the absolute line in both absolute points

A *centre* of a conic in pseudo-Euclidean plane is a pole of the absolute line with the respect to the given conic ([3]).

It is well-known that a conic is uniquely determined by five points or five tangent lines. An *order pencil of conics* is determined by four points

called the *base points* of the order pencil. A *class pencil of conics* is determined by four tangent lines called the *base lines* of the class pencil.

A *curve of centres* is locus of all centres of all conics in a given pencil.

2. CURVE OF CENTRES IN ORDER PENCIL OF CONICS

In the affine model of the pseudo-Euclidean plane (further in text: PE-plane) where the coordinates of the points are defined by

$$x = \frac{x_1}{x_0}, y = \frac{x_2}{x_0} \quad (1)$$

the absolute line f is determined by the equation $x_0 = 0$ and the absolute points F_1, F_2 by the coordinates $(0, 1, \pm 1)$.

Further on, by choosing a basic coordinate simplex in the projective plane P^2 every conic c can be represented by the homogeneous equation of the form

$$a_{00}x_0^2 + a_{11}x_1^2 + a_{22}x_2^2 + 2a_{01}x_0x_1 + 2a_{02}x_0x_2 + 2a_{12}x_1x_2 = 0 \quad (2)$$

and in the affine coordinates by

$$a_{00} + a_{11}x^2 + a_{22}y^2 + 2a_{01}x + 2a_{02}y + 2a_{12}xy = 0 \quad (3)$$

Knowing that the absolute line f has the equation $x_0 = 0$, it is clear from equation (2) that the conic c intersects the absolute line f in the points

coordinates of which satisfy the equality

$$a_{11}x_1^2 + a_{22}x_2^2 + 2a_{12}x_1x_2 = 0 \quad (4)$$

Some short calculations lead to the conclusions:

- c is a hyperbola iff $a_{12}^2 - a_{11}a_{22} > 0$.
 - c is a parabola iff $a_{12}^2 - a_{11}a_{22} = 0$.
 - c is an ellipse iff $a_{12}^2 - a_{11}a_{22} < 0$.
 - c is a special hyperbola iff $a_{11} + a_{22} + 2a_{12} = 0$ or $a_{11} + a_{22} - 2a_{12} = 0$.
 - c is a special parabola iff $a_{11} = a_{22} = -a_{12}$ or $a_{11} = a_{22} = a_{12}$.
 - c is a circle iff $a_{12} = 0$ and $a_{11} = -a_{22}$.
- as shown in [2].

Theorem 1 *Curve of centres of order pencil of conics in PE-plane is a conic.*

Proof:

This is a generalization of the proof given in [1].

Let two conics be given by equations $c_1(x, y) = 0$ and $c_2(x, y) = 0$ where

$$c_1(x, y) = a_{00} + a_{11}x^2 + a_{22}y^2 + 2a_{01}x + 2a_{02}y + 2a_{12}xy \quad (5)$$

and

$$c_2(x, y) = b_{00} + b_{11}x^2 + b_{22}y^2 + 2b_{01}x + 2b_{02}y + 2b_{12}xy \quad (6)$$

Those two conics intersect in four points. If those four points are taken as the base points of the order pencil of conics, then those two conics are called *base conics*. The equation of conic pencil defined by conics c_1 and c_2 as base conics of the pencil is obtained as linear combination of c_1 and c_2 :

$$c_1(x, y) + \lambda c_2(x, y) = 0 \quad (7)$$

where every conic of the pencil is uniquely determined by a $\lambda \in \mathbf{R} \cup \infty$.

It is now easy to calculate the coordinates of a centre of a conic determined by a λ ([3]):

$$x_S = \frac{1}{K}(a_{02}a_{12} - a_{01}a_{22} - \lambda a_{22}b_{01} + \lambda a_{12}b_{02} +$$

$$+ \lambda a_{02}b_{12} + \lambda^2 b_{02}b_{12} - \lambda a_{01}b_{22} - \lambda^2 b_{01}b_{22})$$

and

$$y_S = \frac{1}{K}(-a_{02}a_{11} + a_{01}a_{12} + \lambda a_{12}b_{01} - \lambda a_{11}b_{02} -$$

$$- \lambda a_{02}b_{11} - \lambda^2 b_{02}b_{11} + \lambda a_{01}b_{12} + \lambda^2 b_{01}b_{12})$$

where

$$K = -a_{12}^2 + a_{11}a_{22} + \lambda a_{22}b_{11} - 2\lambda a_{12}b_{12} - \lambda^2 b_{12}^2 + \lambda a_{11}b_{22} + \lambda^2 b_{11}b_{22}$$

Eliminating λ , the equation of the curve of centres is obtained:

$$-a_{02}b_{01} + a_{01}b_{02} - a_{12}b_{01}x + a_{11}b_{02}x - \quad (8)$$

$$-a_{02}b_{11}x + a_{01}b_{12}x - a_{12}b_{11}x^2 + a_{11}b_{12}x^2 -$$

$$-a_{22}b_{01}y + a_{12}b_{02}y - a_{02}b_{12}y + a_{01}b_{22}y -$$

$$-a_{22}b_{11}xy + a_{11}b_{22}xy - a_{22}b_{12}y^2 + a_{12}b_{22}y^2 = 0$$

It is proved that the curve of centres is a conic.

One example of curve of centres k for base conics c_1 and c_2 is shown in Figure 1.

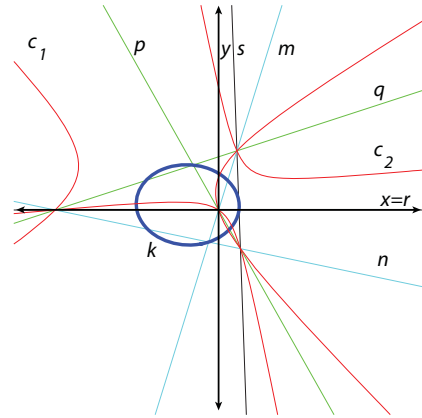


Figure 1: Centres

Four base points of order pencil of conics define six lines which in pairs make three degenerate conics of the pencil: m and n ; p and q ; r and

s. Intersection of each pair of lines is the centre for each degenerate conic. It is easy to calculate that the curve of centres passes through all three centres of degenerate conics of the pencil. This is shown in Figure 1.

Corollary 1 *Curve of centres of order pencil of conics in PE-plane is a hyperbola, ellipse or parabola if and only if two parabolas of the pencil are two real and distinct, two conjugate - imaginary or one double real respectively.*

Proof:

The curve of centres (8) has two intersections with the absolute line f . Knowing that the absolute line has the equation $x_0 = 0$ and writing (8) in homogeneous coordinates, it is simple to obtain the coordinates of these intersections:

$$T_{1,2}(0, 1, \frac{a_{22}b_{11}-a_{11}b_{22}\pm\sqrt{D_1}}{2(-a_{22}b_{12}+a_{12}b_{22})})$$

where

$$D_1 = a_{22}^2b_{11}^2 - 4a_{12}a_{22}b_{11}b_{12} + 4a_{11}a_{22}b_{12}^2 + \quad (9)$$

$$+ 4a_{12}^2b_{11}b_{22} - 2a_{11}a_{22}b_{11}b_{22} - 4a_{11}a_{12}b_{12}b_{22} + a_{11}^2b_{22}^2$$

Obviously, the curve of centres is:

- a hyperbola iff $D_1 > 0$
- an ellipse iff $D_1 < 0$
- a parabola iff $D_1 = 0$.

On the other hand, calculating λ that determines two parabolas of the order pencil (7), it is easy to obtain:

$$\lambda_{1,2} = \frac{a_{22}b_{11}-2a_{12}b_{12}+a_{11}b_{22}\pm\sqrt{D_2}}{2(b_{12}^2-b_{11}b_{22})}$$

where

$$D_2 = a_{22}^2b_{11}^2 - 4a_{12}a_{22}b_{11}b_{12} + 4a_{11}a_{22}b_{12}^2 + \quad (10)$$

$$+ 4a_{12}^2b_{11}b_{22} - 2a_{11}a_{22}b_{11}b_{22} - 4a_{11}a_{12}b_{12}b_{22} + a_{11}^2b_{22}^2$$

This means that:

- the pencil has two real and distinct parabolas iff $D_2 > 0$
- the pencil has two conjugate-imaginary parabolas iff $D_2 < 0$

- the pencil has one double real parabola iff $D_2 = 0$.

Since $D_1 = D_2$, Corollary 1 is proved. Note: In the definition (7) for the order pencil, it is taken $\lambda \in \mathbf{R} \cup \infty$. So in case of two conjugate-imaginary parabolas, it might be better to say: there are no real parabolas.

Corollary 2 *Curve of centres of order pencil of conics in PE-plane is a special hyperbola if and only if one of the parabolas of the pencil is a special parabola.*

Proof:

Let one of the base conics (say, $c_1(x,y)$) of the order pencil (7) be a special parabola. Knowing that a conic (2) (or (3)) is a special parabola iff $a_{11} = a_{22} = \pm a_{12}$, it is easy to calculate that the curve of centres is then:

$$\begin{aligned} & -a_{02}b_{01} + a_{01}b_{02} - a_{11}b_{01}x + a_{11}b_{02}x - \\ & a_{02}b_{11}x + a_{01}b_{12}x - a_{11}b_{11}x^2 + a_{11}b_{12}x^2 - \\ & a_{11}b_{01}y + a_{11}b_{02}y - a_{02}b_{12}y + a_{01}b_{22}y - \\ & a_{11}b_{11}xy + a_{11}b_{22}xy - a_{11}b_{12}y^2 + a_{11}b_{22}y^2 = 0 \end{aligned}$$

in case $a_{11} = a_{22} = a_{12}$

and

$$\begin{aligned} & -a_{02}b_{01} + a_{01}b_{02} + a_{11}b_{01}x + a_{11}b_{02}x - \\ & a_{02}b_{11}x + a_{01}b_{12}x + a_{11}b_{11}x^2 + a_{11}b_{12}x^2 - \\ & a_{11}b_{01}y - a_{11}b_{02}y - a_{02}b_{12}y + a_{01}b_{22}y - \\ & a_{11}b_{11}xy + a_{11}b_{22}xy - a_{11}b_{12}y^2 - a_{11}b_{22}y^2 = 0 \end{aligned}$$

in case $a_{11} = a_{22} = -a_{12}$.

The coefficients of x^2 , y^2 and xy are such that it is clear that the curve of centres is a special hyperbola.

It remains to prove that if the curve of centres is a special hyperbola, then one parabola of the pencil is a special parabola. In order to have a special hyperbola as the curve of centres (8), the following condition must be satisfied:

$$-a_{12}b_{11} + a_{11}b_{12} + (-a_{22}b_{12} + a_{12}b_{22}) + (-a_{22}b_{11} + a_{11}b_{22}) = 0$$

or

$$-a_{12}b_{11} + a_{11}b_{12} + (-a_{22}b_{12} + a_{12}b_{22}) - (-a_{22}b_{11} + a_{11}b_{22}) = 0$$

Let us take the first one (and the second one may be proved in the same way). In that case, the coefficient a_{11} is:

$$a_{11} = \frac{a_{12}b_{11} + a_{22}b_{11} + a_{22}b_{12} - a_{12}b_{22}}{b_{12} + b_{22}}$$

In that case, the pencil (7) has the equation:

$$a_{00} + \lambda b_{00} + 2a_{01}x + 2\lambda b_{01}x + \lambda b_{11}x^2 + \frac{a_{12}b_{11}x^2}{b_{12} + b_{22}} + \frac{a_{22}b_{11}x^2}{b_{12} + b_{22}} + \frac{a_{22}b_{12}x^2}{b_{12} + b_{22}} - \frac{a_{12}b_{22}x^2}{b_{12} + b_{22}} + 2a_{02}y + 2\lambda b_{02}y + 2a_{12}xy + 2\lambda b_{12}xy + a_{22}y^2 + \lambda b_{22}y^2 = 0$$

Calculating λ that determines parabolas of the pencil, it is obtained:

$$\lambda_1 = \frac{-a_{12} - a_{22}}{b_{12} + b_{22}}$$

and

$$\lambda_2 = \frac{-a_{22}b_{11} + a_{12}b_{12} - a_{22}b_{12} + a_{12}b_{22}}{-b_{12}^2 + b_{11}b_{22}}.$$

For λ_1 , the parabola touches the absolute line f in the point with the coordinates $(0, 1, 1)$ and that is the absolute point, so it is clear that this parabola is special parabola and Corollary 2 is proved. In the same way the following two corollaries may be proved:

Corollary 3 *Curve of centres of order pencil of conics in PE-plane is a circle if and only if both parabolas of the pencil are special parabolas.*

Corollary 4 *Curve of centres of order pencil of conics in PE-plane is a special parabola if and only if the pencil has one double real parabola and it is a special parabola.*

3. CURVE OF CENTRES IN CLASS PENCIL OF CONICS

Theorem 2 *Curve of centres of class pencil of conics in PE-plane is a line.*

Proof:

Let two base conics be given by equations in line coordinates $c_1(u, v) = 0$ and $c_2(u, v) = 0$ where

$$c_1(x, y) = c_{00} + c_{11}u^2 + c_{22}v^2 + 2c_{01}u + 2c_{02}v + 2c_{12}uv \quad (11)$$

and

$$c_2(x, y) = d_{00} + d_{11}u^2 + d_{22}v^2 + 2d_{01}u + 2d_{02}v + 2d_{12}uv \quad (12)$$

$$+ 2d_{02}v + 2d_{12}uv$$

and let their linear combination define a class pencil of conics:

$$c_1(u, v) + \lambda c_2(u, v) = 0 \quad (13)$$

where every conic of the pencil is uniquely determined by a $\lambda \in \mathbf{R} \cup \infty$.

It is now easy to calculate the coordinates of a centre of a conic determined by a λ :

$$x_S = \frac{c_{01} + \lambda d_{01}}{c_{00} + \lambda d_{00}}$$

and

$$y_S = \frac{c_{02} + \lambda d_{02}}{c_{00} + \lambda d_{00}}$$

Eliminating λ , the equation of the curve of centres is obtained:

$$c_{02}d_{01} - c_{01}d_{02} - c_{02}d_{00}x + c_{00}d_{02}x + \quad (14)$$

$$+ c_{01}d_{00}y - c_{00}d_{01}y = 0$$

which is clearly a line and Theorem 2 is proved.

4. CONCLUSIONS

In this paper it is proved that in pseudo-Euclidean plane the curve of centres of an order conic pencil is a conic. It is proved that the curve of centres can be any type of PE-conic and it is explained how that depends on the type of conics which are contained in the given pencil, mainly on the type of parabolas in the pencil. It is also proved that the curve of centres of a class conic pencil is a line.

REFERENCES

- [1] M. Katić Žlepalo, Curves of Foci, Vertices, Centres and Directrices of Pencils of Conics in the pseudo-Euclidean Plane, PhD Thesis, Zagreb, 2014.
- [2] N. Kovačević and E. Jurkin, Circular Cubics and Quartics in Pseudo-Euclidean Plane Obtained by Inversion, *Math. Pann.* 22 (2) (2011), 199-218

- [3] D. Palman, Projektivna geometrija, Školska knjiga, Zagreb, 1984.
- [4] N. V. Reveruk, N. V.: Krivie vtorogo porjadka v psevdoevklidovoi geometrii, *Uchenye zazapiski Moskovskogo pedagogicheskogo instituta* 253, Moscow, 1969.
- [5] H. Sachs, Ebene Isotrope Geometrie, Wieweg, Braunschweig-Wiesbaden, 1987

ABOUT THE AUTHORS

1. Mirela Katić Žlepalo, Ph. D., is Lecturer at the Polytechnic of Zagreb, Department for Civil Engineering and secretary of the Croatian Society for Geometry and Graphics. She can be reached by e-mail: mkatic@tvz.hr.

DESCRIPTIVE GEOMETRY EDUCATION BY USING MULTIMEDIA TOOLS

Aleksandar A. ČUČAKOVIĆ, Nataša K. TEOFILOVIĆ and Biljana S. JOVIĆ

The University of Belgrade, Serbia

ABSTRACT: This paper proposed integration of multimedia and 3D animation tools in the descriptive geometrical education. This transdisciplinary and hybrid visual dynamic educational tool is aimed toward students of technical and applied arts. Computer technology in the function of geometry learning tools offers new and fascinating possibilities. Students and teachers can explore the most diverse theoretical and practical problems with the aim of understanding the dynamic and complex spatial relationships.

The multimedia consists of short animated forms supported by concise textual explanations. Sixteen integrated animated short forms are five minutes duration in average. DVD title is "Geometric education using the principles and tools of 3D animation". Following descriptive geometrical areas are covered: 1. Platonic solids: cube, tetrahedron, octahedron, dodecahedron and icosahedrons. 2. Ruled surface: conoid, rotational hyperboloid, helicoid and hyperbolic paraboloid. 3. The revolved surfaces: torus. 4. Mutual intersection: conic sections, cone and cylinder, sphere and cylinder and two half-cylinder. 5. Experimental design (freeform): generating a surface with the two profiles as guidelines, generating free form using lattice deformers and generate free-form by the duplicating along curves tool. The aim of this learning method is to simplify the perception of geometrical forms, the process of their constructions, and to inspire the users to freely and without fear further explore this complex subject.

Geometrical education using 3D animation principles and tools represents a new methodology approach and its final result is presented on new way, as multimedia tool. Its originality is based on interdisciplinary approach and use of new technologies. The method derived from the overlapping of multiple disciplines such as descriptive geometry, computer animation, structural systems and programming. The results are multilayered and are the base for further scientific research and new development in practice.

Aspects and research results presented in this paper are: the final application in education, production methodology and the basic structure of multimedia learning aid.

Keywords: Descriptive geometry, education, computer animation, multimedia

1. INTRODUCTION

Perception and shaping space is the most essential process in the education of art and engineering students, especially at the technical and technological group. Geometrical education includes spatial understanding and mental visualization of space structures. This paper is based on the hypothesis that the application of CG computer animation in education could be used as additional education tool in Descriptive geometry subject. This work is research in the field of application of methodological innovation in the area of space

geometry and computer animations. The geometrical education involves the geometry in the plane and space geometry. The applicability of dynamic 3D geometry in education provides improvement of spatial ability [6]. It is pedagogical stimuli for users in terms of encouragement for the further geometry exploration [5]. The proposed form of educational environments is a new potential for the study of geometry.

2. THE PERCEPTION IN EDUCATION

Of all the sensory sensations inherent to physiology of the humans more than 70%

belong to the field of visual perception or the sense of vision. For the human vision, without doubt, can be argued that it is most important sense for the human race. A special feature of visual perception is the capacity to grasp the three-dimensional space [4].

The comprehension ability of three-dimensional space means to estimate the relative distances between objects as well as the inner sense of form shape, i.e. the third dimension of space [7]. From this reason the animations on multimedia DVD are presented as a short animations with a concise text explanations (Figure 1.).

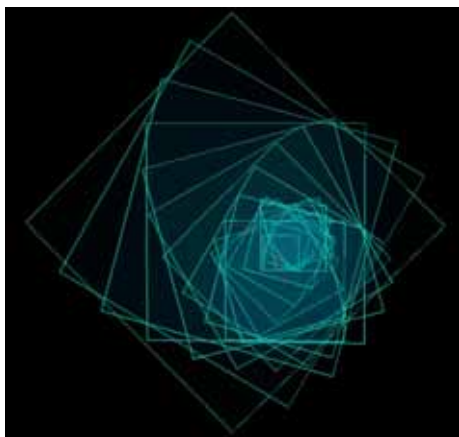


Figure 1: Cover page of the multimedia DVD: Geometrical education by using principles and tools of 3D animation

3. DISCUSSION

Application of CG computer animation as additional tool in education helps students to solve geometrical problems and to perceive the geometry in 3D space [2].

To analyze the application of the proposed method in practice in the education, multimedia material is presented as a supplementary tool in teaching Descriptive geometry [3].

The proposed method is presented to 40 students on first and second year of studies on the Descriptive geometry subject in regular

classes at Department of Landscape architecture and horticulture, Faculty of Forestry, University of Belgrade.

On the first year of studies students learn Geometry by using paper and pen as well as multimedia DVD. Students also made some geometrical models using paper in conventional way (Figure 2.).



Figure 2: Students of Landscape architecture made geometrical models

On the second year of studies students of the Landscape architecture learn geometry using computers and also using DVD multimedia teaching materials.

After two years in geometrical education using different type of educational material students are kindly ask to answer our mixed type of questionnaire.

3.1 Analysis of the results in the practice

Results were analyzed after the Course of Descriptive geometry and Landscape - architecture graphics through questionnaire. Questionnaire had the following structure divided in three parts:

1. Whether students were previously known or familiar with the use of multimedia in education;
2. Students compared multimedia educational tool to conventional, printed materials that are already in use;
3. Students gave analysis of presented geometric areas and named the areas that they feel the need to be processed in this way and were not shown on multimedia DVD.

The responses to the questionnaire were as follows:

1. Use of video tutorials available over the Internet in various fields is not unknown to students. As for the specific application, in this field of Descriptive geometry, they perceive it as an innovation.
2. The present material was very well accepted. Material presented on multimedia DVD kept student's attention. They even expressed a desire to DVD has sound and narration. The largest percentage of them see this as an excellent supplementary material tool in the education. In students opinion advantages of printed material are: possibility of depth analysis and great quantity of information and the ability to take notes. The main advantage of using multimedia DVD is that students could see the spatial forms and understand the problem in 3D space. In particular, the positive evaluation refers to possibility to follow gradual construction of objects, both form and constructions process in several projections which in classic textbooks is not possible.
3. In this context they expressed a desire to use multimedia DVD as a learning tool in other geometrical areas such as perspective which is not produced in multimedia form yet.

4. CONCLUSIONS

New instruments allow students, teachers, artists, researchers, engineers, designers, etc.. empowerment in terms of strengthening the contribution in their field of work. Results are richer, it is possible to learn more than you can learn only from books and often simpler for using than working with paper and pencil.

Students and teachers can explore the most diverse theoretical and practical problems with the aim of understanding the dynamic and complex spatial relationships [8].

These capabilities offer a new way of conversation and communication between

teachers and students which at conventional ways of teaching are not obvious or possible. The use of multimedia computer animation in the teaching related to the geometrical education improves and greatly speeds up explanations of teachers intentions [1].

Experience in the use of multimedia computer animation in the learning process demonstrates significant progress in the perception of huge possibilities working with each model.

The analysis of the questionnaire shows that the innovation positively evaluates the use of multimedia in its dynamic dimension as animation have feature which highlights process in constructing geometric shapes and solving geometric problems.

Dynamic geometry education achieves much higher insight into the actual structure and construction, where through the movement directly and experimentally we learn about the changes in the construction of the structure [9].

The original contribution of this paper is in the implementation of multiple disciplines, and this interdisciplinary hybrid approach, overlapped several disciplines such as architecture, descriptive geometry, computer animation and programming.

ACKNOWLEDGMENT

Authors are supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia, Project No. TP 36008

REFERENCES

- [1] Čučaković and B.Jović. Constructive Geometry Education by Contemporary Technologies. *SAJ_2011_3_Serbian Architectural Journal*, original scientific article, UDK 514.18:62 ID 184977420 pp. 164-183, Belgrade, Serbia, 2011.
- [2] Čučaković, M.Nestorović and B.Jović. Contemporary principles of geometrical modeling in education. *Abstracts - 2nd Croatian Conference of*

- Geometry and Graphics Scientific-Professional Colloquium of CSGG(Šibenik, Croatia, September 5-9), pp. 10-11, Šibenik, Croatia, 2010.*
- [3] Čučaković and B.Jović. Optional Course Engineering Graphics On Department For Landscaping Architecture At The Faculty Of Forestry, University Of Belgrade. (*International Conference SUNGIG moNGeometry, June 24-27, Belgrade, Serbia, 2010.*)
 - [4] Jović. Geometrijska edukacija na polju vizuelizacije i eksperimentalnog dizajna primenom virtuelnih tehnologija. Doktorska disertacija, Arhitektonski fakultet Univerziteta u Beogradu, Beograd, Srbija, 2012.
 - [5] H. Kaufmann. Geometry education with augmented reality. Dissertation, Technology University of Vienna, Austria, Vienna, 2004.
 - [6] U.H. Kortenkamp. Foundation of Dynamic Geometry. PhD Dissertation. Swiss Federal Institute of Technology, Zurich, Switzerland, 1999.
 - [7] H. Stachel. "[*What is Descriptive Geometry for?*](#)" In DSG-CK Dresden Symposium Geometrie: konstruktiv & kinematisch (*Dresden, Germany, Feb. 27 - March 1, 2003.*), (ISBN 3-86005-394-9): pp. 327-336. TU Dresden, 2003.
 - [8] N. Teofilović. Umetnost pokreta u prostoru praznine (tehnologija i praksa virtuelnih karaktera). Arhitektonski fakultet Univerziteta u Beogradu, Srbija, Beograd, 2011.
 - [9] N. Teofilović, 1:1 (3D karakter animacija i instalacija). Doktorski rad. Univerzitet umetnosti u Beogradu, Interdisciplinarne doktorske studije, Grupa za digitalnu umetnost, Beograd, 2010.
- Math, Phys, and Descriptive Geometry, University of Belgrade, Serbia. Fields of interest: descriptive geometry, perspective projections, projective geometry, computational geometry, geometry of surfaces, tessellation, geometric education, architecture, etc.
E-mail: cucak@grf.bg.ac.rs
- 2. Nataša Teofilović
PhD, digital artist, MA architect, lecturer Faculty of Architecture, University of Belgrade, Serbia. Fields of interest: digital art, new media art, 3D character animation, 3D character animation education, etc.
E-mail: nnwt.nat@gmail.com
 - 3. Biljana Jović
PhD in architecture, MSc in architecture, landscape architect, Docent, Faculty of Forestry, Department of landscape architecture and horticulture, University of Belgrade, Serbia. Fields of interest: descriptive geometry, anaglyph projections, projective geometry, computational geometry, multimedia, geometrical education, architecture, landscape architecture, geometrical art etc.
E-mail: biljana.jovic@sfb.bg.ac.rs

ABOUT THE AUTHORS

- 1. Aleksandar Čučaković
PhD, architect, associated professor, Faculty of Civil Engineering – Institute of

DESIGN AND APPLICATION STUDIES FOR A CUPOLA FORMING ORBITAL ARRANGEMENT OF MIURA-ORI BASIC UNITS

Martin BAREJ¹, Susanne HOFFMANN², Christian KNIEF¹,
Martin TRAUTZ² and Burkhard CORVES¹

¹Dept. of Mechanism Theory and Dynamics of Machines, RWTH Aachen University, Germany

²Chair of Structures and Structural Design, RWTH Aachen University, Germany

Deployable folded plate structures are a promising design principle for technical applications. Their two main abilities: providing stiffness in the folded or partially folded state and changing their shape between a folded and an unfolded state make them predestinated especially for self-supporting convertible structures in architecture.

For reasons of cost effectiveness and assembly, structures based on repeating folding patterns like those from tessellation origami are the preferred source of inspiration. The probably most studied folding pattern is Miura-Ori. Its variations reach from fully regular patterns consisting of only one type of elements to completely irregular ones. Derivatives can cover tapered areas instead of rectangular ones or form arcs instead of remaining planar. Also characteristics like developability or flat-foldability can be changed, but some characteristics are always constant. Firstly there are always four folds meeting at each vertex, whereby there are one mountain fold and three valley folds or the other way round. Secondly in the unfolded state a mesh of folds can be identified, where lengthwise the folds change their direction from mountain to valley and back but crosswise they do not. Neglecting this second aspect a pattern can be created which consists of an orbital arrangement of Miura-Ori basic units. It can fold from a planar pattern to a sector whereby the radially outer segments coil up, as shown in Figure 1.



Figure 1: Folding Sequence of the 4-Segmented Pattern

This paper presents variations of this pattern, whereby different configurations as well as changes of dimensional parameters are taken into account. The different variations are characterized by specifications like their numbers of elements and folds, the covered area or the folding sequence.

The projection of possible applications includes concepts using the unfolded but also concepts using the folded or partially folded state.

Keywords: Tessellation Origami, Variations of Miura-Ori, Orbital Arrangement of Miura-Ori Basic Units, Application Studies, Foldable Structures in Architecture

1. DEPLOYABLE FOLDED PLATE STRUCTURES

Seemingly architecture consists of immobile buildings. However, even the most inconspicuous building comprises convertible components such as doors and windows. And on closer inspection, the field of architecture consists of a large variety of transformable structures like mobile pavilions, adaptive facades or even convertible roofs and retractable bridges. Because the demand towards sustainability increases, adaptive building parts as well as buildings for temporary use will be more required so that the importance of transformable architecture is expected to increase.

Foldable structures are adaptable on the one hand and form a self-supporting structure on the other hand. Due to these two key properties – strengthening and deployability – foldable structures are predestined for transformable constructions in architecture and other technical applications. Thereby the most promising foldable structures are based on the so called Origami Tessellations because these consist of repeating patterns [1] and for that reason only a limited number of modules are necessary.

In spite of these evident advantages realized folded plate structures are rare. In architecture complex foldable structures are found almost exclusively in subsystems e.g. adaptable façade elements or smaller object architecture as shown in Figures 2 and 3.



Figure 2: Al Bahr Towers,
Aedas



Figure 3: Canary Wharf Kiosk,
Make Architects [2]

Large-scaled convertible Origami-based folded plate structures are non-existent. This is probably due to a disproportional rising of challenges for the production process and specially designed connecting details with increasing complexity of the folding pattern.

2. MIURA-ORI

The first step to reduce the complexity of a foldable structure is the use of a manageable folding pattern. A widespread solution is Miura-Ori shown in Figure 4.

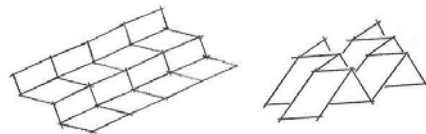


Figure 4: Miura-Ori

Originally, this pattern was developed to compact large membrane structures for transportation to space [3]. But its regular arrangement, the fact that it also works with rigid plates and its constrained motion make it predestinated also for applications in architecture.

2.1 Variations of Miura-Ori

In the original Miura-Ori pattern four similar parallelograms are connected by their edges in a way that the four folding axes meet in one node. Such a basic unit is deployable with one

degree of freedom. Further it is developable and flat-foldable.

A regular array of these basic units is characterized by parallel lengthwise folds with alternating folding directions (mountain/valley) and crosswise folds with a zigzag course which are either mountain or valley folds over the complete width of the array. In this assembly all nodes and their folding angles are identical.

From this base, variations of the pattern can be derived as studied in [4] and [5]. Changes in the dimensions or angles of parallelograms affect the dimensions of the folded and unfolded structure but do not influence the global appearance. Replacing parallelograms by trapezes, rectangles or general quadrangles can lead to surfaces folding into a curve instead of contracting in a plane (Figure 5).

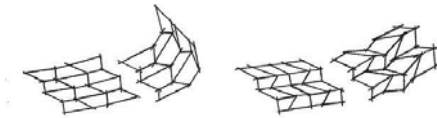


Figure 5: Curve-Forming Variations
(according to [4])

If the lengthwise folds intersect each other instead of being parallel a tapered version of the pattern appears. Allowing changes in the direction of these folds can lead to irregular patterns. (Figure 1)

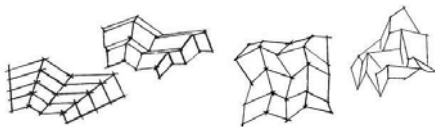


Figure 6: Tapered (according to [4])
and Irregular Variations

The pattern even works when the plate's angles at a node are not equal to 360° . In these cases the resultant structure is not flat foldable or not developable.

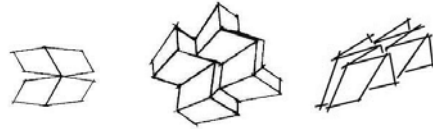


Figure 7: Non-Developable Variations
(according to [4])

2.2 Cupola Forming Orbital Arrangement

While planar or curve forming variations of the Miura pattern can be build up using only one type of elements, the tapered arrangement requires bigger elements for every orbital row. An attempt to assemble a tapered pattern with only one type of elements is to arrange Miura-Ori basic units of symmetric trapezes as depicted in Figure 8.

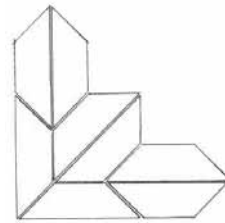


Figure 8: Crease Pattern

In this crease pattern the long bases of the trapezes are always connected to other long bases, but short bases and legs need to be connectable among each other. Hence symmetry is not sufficient as criterion but the trapezes additionally need to be triamonds.

Like other tapered patterns the presented one requires a radial cut for folding. Starting from an un-folded or developed state its motion sequence (Figure 9 A-C) is a simultaneous combination of an angular folding around a vertical axis and a coiling up around a horizontal one. Assuming a negligible thickness the final state is flat folded.

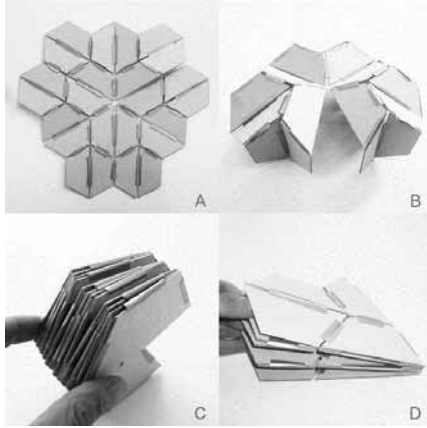


Figure 9: Motion Sequence

In general three states are regarded for Origami structures:

- A: un-folded or developed,
- B: partially folded and
- C: folded ideally flat-folded.

For practical reasons like space requirements for transport or storage an additional state may be necessary:

- D: folded parts (alternative way)

For this state the structure is disassembled into sub-structures. The folding sequence of the sub-structure is based on the chosen crease pattern, but it is not necessarily using all folds or it may be using valley folds as mountain folds or the other way round.

3. DESIGN STUDY

Based on the crease pattern described in Figure 8 different designs are possible. These can be reached by changing geometrical parameters of single elements or by adding elements in angular direction and thereby extending the structure.

3.1 Parameterization and Extension

Since the crease pattern is a tessellation of triangles for every element three of its edge lengths are identical and the fourth one is dependent on them. Hence there are only two

independent parameters for a trapeze as depicted in Figure 10:

- one angle and
- one edge length.

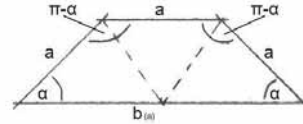


Figure 10: Parameterization of Trapeze

Depending on the chosen angle α the structure can be extended in angular direction up to the number of $360^\circ/\alpha$ elements in the inner row. Further extension would cause overlapping. In radial direction the unfolded tessellation is endlessly extendable, but a continuation of the pattern further as depicted in Figure 8 can lead to intersections of elements in the folded state.

3.2 Effects on Geometry

The first and very basic step in the analysis of effects of variations is changing the parameters of elements. Varying the edge length influences characteristics like:

- area or diameter in the unfolded state
- base area, diameter or enclosed space in the partially folded state
- area of the cross section in the flat folded state

The same characteristics also can be influenced by changes of the trapeze's angle. Additionally reducing the trapeze's angle makes the cut angle rise. Assuming that in the unfolded state the structure should form a closed surface the open space in the cut angle must be filled by angular extension. So there are influences on:

- cut angle in the unfolded state
- number of elements (if extended)
- number of folds (if extended)

In a design study characteristics of the depicted variations are revealed and their respective benefits in terms of suitability for a given task are verified. Furthermore the results of the

study give a perspective which variations could match other use cases. In the following designs with three up to ten segments are studied. Their trapeze angles and the used representing symbols are given in Table 1.

Table 1: Variations of Crease Pattern

Number of Segments	Trapeze's Angle α	Symbol
3	60°	◆
4	45°	◇
5	36°	▲
6	30°	△
7	25,7°	■
8	22,5°	□
9	20°	●
10	18°	○

To enable a comparison first a uniform task needs to be set and all variations have to be adapted to fulfill it. Assuming that in the partially folded state the given pattern might be used as a temporary building the chosen task is to offer a useable space with a ground area of at least 12m² and a height of at least 2.2m (Figure 11).

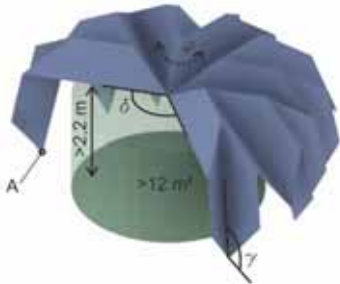


Figure 11: Given Task

This task is regarded as fulfilled when a cylinder with the given area and height can be integrated inside as shown in Figure 11. The fact that a part of the area is not covered by a ceiling is neglected.

To meet the requirements there is only one free geometrical parameter (the edge length), but since the regarded state ‘partially folded’ is not exactly defined yet also the folding angle φ is still free to choose. Therefore as a second condition, the opening angle in the cut δ has to accommodate an integral number of additional segments with the same folding angle as the partially folded structure.

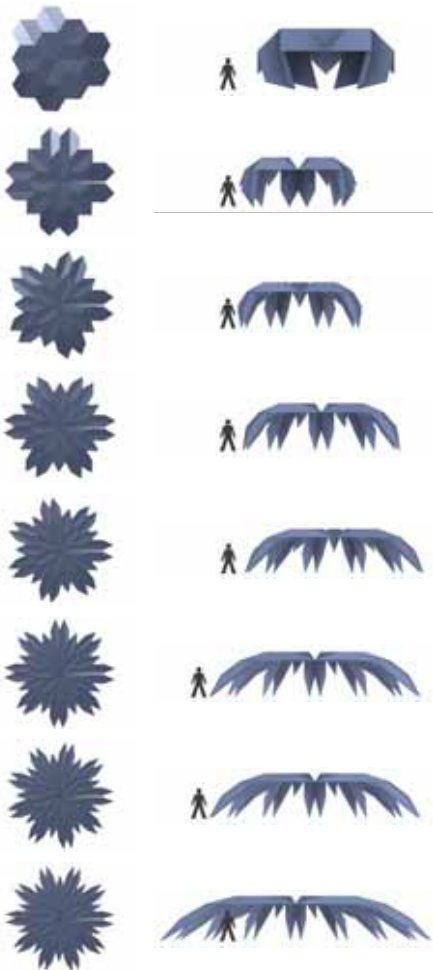


Figure 12: Extensions of Crease Pattern

Since this condition still allows several variations it is defined that the 3-segmented pattern has to accommodate one additional segment. With any higher number of segments in the structure also the number of segments for the gap is raised by one.

The chosen set of conditions leads to the patterns and structures depicted in Figure 12. Their folding angles and opening angles are compiled in Table 2. It can be seen that for patterns with high numbers of segments the values of both state describing angles (γ , δ) hardly differ from pattern to pattern in contrast to differences between patterns with small numbers of segments. This behavior already could be observed at the trapeze's angles α .

Table 2: Angles in the Partially Folded State

Pattern	Folding Angle φ	Opening Angle δ	Wall Angle γ
◆	109,5°	90°	70,5°
◇	90°	120°	101,5°
▲	81,2°	135°	118,3°
△	76,3°	144°	129°
■	73,2°	150°	136,3°
□	71,1°	154,3°	141,9°
●	69,5°	157,5°	146,1°
○	68,4°	160°	149,5°

Regarding the angles the first outcomes are that patterns with high numbers of segments leave a bigger sector uncovered but for their construction hinges with a smaller motion range can be used than for example for the 4-segmented pattern with its folding angle of 90°.

Another characteristic angle is the one between the ground and wall in the partially folded state. As shown in Table 2 the walls of the 4-segmented pattern stop close to the vertical position. Patterns with more segments display larger angles. The 3-segmented pattern shows an anomaly. Here the walls pass the rights angle and do not stop before reaching about 70°. At first glance this only means that

for the chosen conditions the area of the floor is smaller than the area of the ceiling which not necessarily is a problem. But there are also other consequences on the folding sequence in use and on the structures' stability.

As mentioned already in the introduction one key aspect for the choice of a pattern is its complexity. Under this term different characteristics are concentrated, amongst others the numbers of elements and folds (Figure 13). The fewer elements and folds are necessary the lower is the complexity and the more beneficial is the structure with regard to assembly or failure risk. Since the patterns differ only by the number of segments, with any segment the numbers of elements and folds increase equally spaced: +8 elements and +12 folds.

Also the use of different element shapes or different joints would increase complexity however the regarded patterns consist of only one type of elements and only one type joints respectively. In a later design phase detailed engineering may require different joints e.g. caused by adaption to loads or necessary motion ranges.

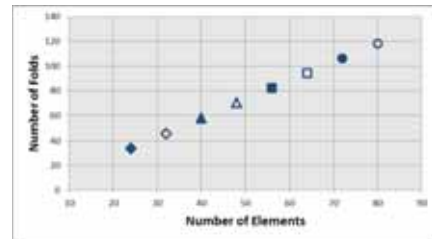


Figure 13: Complexity

Another aspect is the geometry of elements required to fulfill the given task. Large-area, heavy elements complicate the assembly of a structure. Further they are sensitive to deformation and thereby may impede the folding motion. The altitudes of elements are related to the trapezes' angles which directly depend on the patterns' numbers of segments. While between the altitudes of the 3-segmented and the

4-segmented pattern there is a factor of about two, patterns with higher numbers of segments have little differences between their trapezes' angles so their altitudes are more alike. The increase of lengths is related to the partially folded state. Patterns with high numbers of segments reach the required state with lower changes of folding angles. That means that the coiling up of the outer elements is not so advanced and thus the elements need to be longer to achieve the required height of the cupola. The outlier of the three-segmented pattern is due to the fact that here the angle γ between ground and side wall is about 70° and thus the inclination of the wall requires longer elements than at the 4-segmented pattern with an angle of approx. 100° .

According to the results depicted in Figure 14 the 4-, 5- and 6-segmented patterns seem to offer manageable areas and suitable ratios of lengths and altitudes of elements. Detailed consideration may be reasonable, when there are concrete space requirements for transport or storage.

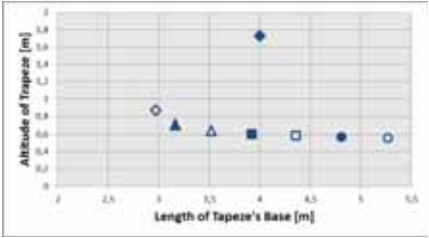


Figure 14: Geometry of Elements

Regarding sustainability a high material efficiency is to aspire. In Figure 15 the offered areas of the structures and the accumulated areas of their elements are given. According to results from Figures 13 and 14 the material usage is low for structures with little numbers of elements. Patterns with more segments require more material but due to their longer elements they also offer more useable space than required by the given task. Therefore they may be beneficial for use cases with higher space

demands. A distinction between covered and open area is neglected again.

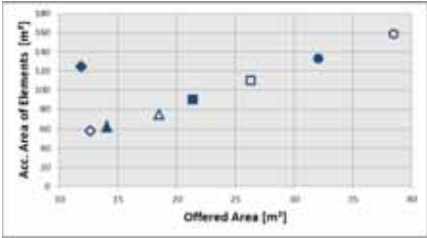


Figure 15: Material Efficiency

For the comparison of motion sequences the path of vertex A (Figure 11) can be regarded. Figure 16 shows its path on a vertical plane rotating around the structure's central vertical axis. The use of this rotating plane reveals the effective radius instead of a projection. The motion paths start at the folding angle of 175° and swing down to the partially folded state defined as use case. The particular path of the 3-segmented structure is caused by its bigger elements and its wall angle $\gamma < 90^\circ$. The bigger elements effect the higher radius at the beginning of the motion while the wall angle makes a downswing below the final value of 2.5 m necessary.

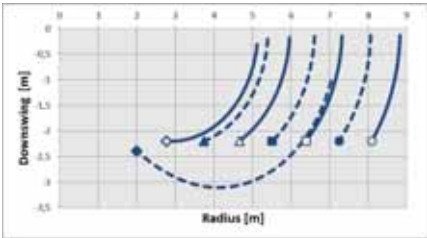


Figure 16: Path of 'A' in Vertical Plane

The topview on the motion of vertex A is given in Figure 17. The representing symbols of the patterns are placed where 'A' stops in use position. The side from which this position is achieved depends on the starting state. The upper paths represent the unfolding from fold-

ing angles equal to 10° to the partially folded state. The lower paths comply with the motion from the unfolded state ($\varphi=175^\circ$).

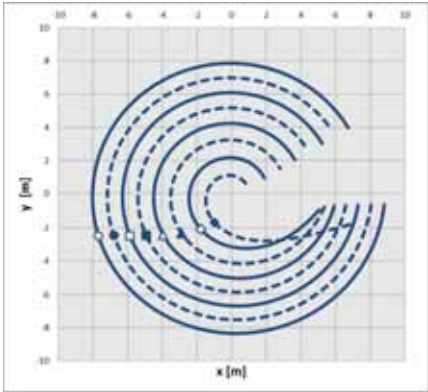


Figure 17: Path of ‘A’ in Horizontal Plane

The paths of patterns with five or more segments differ primarily by their scaling. The 3- and 4-segmented patterns also show different courses. Here the changes of radii are stronger when the motion starts from the unfolded state. That means that the space required to fold the structure into the use state is significantly bigger than in the use state itself. This outcome can be regarded as a disadvantage of these two patterns, but if the structure can be set up by unfolding from the flat folded state this disadvantage carries no weight.

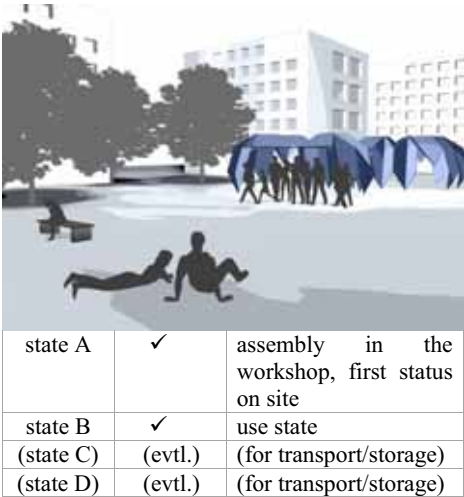
Summing up, the patterns with fewer segments seem to be more suitable for the chosen task. A perhaps unexpected exception is the poor result of the 3-segmented pattern. For this structure the required ratio of height and span width is unfavorable. In use cases requiring a covered area in the unfolded state this structure could show its benefits like low complexity or high compressibility. Patterns with higher numbers of elements offer a wider span and should be taken into account for applications where a bigger useable space is required.

4. APPLICATION STUDIES

The Miura-Ori based cupola forming folding patterns allow a grand variety of possible applications. The field ranges from small-scaled pavilions to large-scaled deployable roofs, from mobile to immobile structures and from using the unfolded state to using a folded state. The chosen examples give a selected overview of the various options these folding patterns can cope with. Furthermore it is indicated which of the states introduced in Figure 9 are relevant for each example.

4.1 Small-Scale Structures

Figure 18 shows the pattern used as a temporarily used small kiosk, for example on a festival or in a park.



state A	✓	assembly in the workshop, first status on site
state B	✓	use state
(state C)	(evtl.)	(for transport/storage)
(state D)	(evtl.)	(for transport/storage)

Figure 18: Kiosk/Pavilion

The kiosk shows the cupola forming foldable structure with its purest characteristics. Out of a flat folded plate structure, it forms a roof structure in its functionality comparable to simple tents, but with much more impression as an attractive design object. Although this structure is heavier and probably more expensive than competitive products it is more robust and on this account it can fulfill more tasks. In

the next example it is utilized as a shelter (Figure 19).

The dimensions are comparable to the first example, but the demands to functionality are higher.



state A	-	-
state B	✓	use state
state C	✓	for transport
(state D)	(evtl.)	(for transport)

Figure 19: Shelter

As a shelter has to provide protection against weathering and protection of privacy it has to be a closed space that the basic patterns do not achieve; extra elements have to be added. Folded state C is now the first option for transport and storage. Even more advantageous would be using state D because of less stacked layers and therefore optimized space-saving.

4.2 Transformable Roof Structure

The challenges are different when the cupola forming folding pattern is used in a larger scale as visualized in Figure 20. The immobile deployable roof structure, e.g. for a stage, fulfills its function as weather protection and supporting structure for technical elements during an event and is a space-saving structure while not in use. In any case it works as an interesting landmark. As shown in Figure 20 the structure can also be used rotated 180°. The structural behavior might be more challenging, but the big advantage of using it this way is that no extra closing elements are necessary.



state A	✓	assembly on site
state B	✓	use state
state C	✓	non-use state
state D	-	

Figure 20: Deployable Roof

4.3 Sunshade Element

Figure 21 shows the foldable structure as a sunshade element.



state A	✓	use state
state B	✓	interstage
state C	✓	non-use state
(state D)	(evtl.)	(for transport/storage)

Figure 21: Sunshade

This works assembled horizontally as well as vertically or in other degrees. Since the folded state A is the use state, large wing-spreads are not possible due to little effective depths. Hence this is an application idea

for innovative small-scaled architectural elements. If not in use, the folded structure (state C) is optimally space-saving. As this structure also consists of only one module it represents a promising alternative to conventional systems. The four shown examples are meant to inspire, the application possibilities are much more versatile.

4.4 Requirements for realization

If a realization of a cupola forming Miura-Ori-based structure is to be designed many requirements have to be fulfilled. Even if these vary from project to project, the following ones are elementary for foldable structures and form at the same time the most challenging ones: storage, transport and assembly concept, concept for repetitive movement, rain and weather protection, security arrangements. Aligned solutions are topic of ongoing interdisciplinary research activities.

5. CONCLUSIONS

The use of deployable folded plate structures provides innovative solutions for technical tasks. Thereby the folding pattern Miura-Ori as the most reduced transformable folding pattern forms the starting point for the investigation of application possibilities in architecture and engineering. The cupola forming Miura-Ori based foldable structure which is analyzed in this paper shows once more the geometric diversity and variability of the folding pattern Miura-Ori. By varying the angles of the trapezes arranged in a staggered manner numerous shapes and thus a wide range of possible applications are generated. With the requirements (for a small pavilion) chosen in the analysis, useful structures ensue from the 4- or 5-segmented pattern. For other applications, e.g. for structures with a wider span more differentiated patterns could be expedient, so that further design studies will be necessary in any case. Generally the complexity of foldable structures increases both with a rising number of elements (and hence resulting rising number of connections) and with more acuted angles. On this account the simplest modular solutions should be aspired.

6. ACKNOWLEDGMENTS

The authors want to thank the Exploratory Research Space at RWTH Aachen University for financing the research project “TransOri – Transferring Origami Principles to Technical Applications”.

REFERENCES

- [1] E. Gjerde. Origami Tessellations: Awe-Inspiring Geometric Designs. A K Peters, Ltd. Natick 2009.
- [2] <http://www.makearchitects.com/projects/cary-wharf-kiosk/> 03.04.2014
- [3] Miura, K.: Method of Packaging and Deployment of Large Membranes in Space. The Institute of Space and Astronautical Science Report No. 618, Tokyo 1985
- [4] Gattas, J. M.; Zhong, Y.: Rigid-Foldable Piecewise Geometries. In: Proceedings of the First Conference Transformables 2013
- [5] Tachi, T.: Generalization of Rigid Foldable Quadrilateral Mesh Origami. In: IASS Symposium, Valencia 2009

ABOUT THE AUTHORS

1. Martin Barej,
scientific assistant at the
Department of Mechanism Theory
and Dynamics of Machines,
RWTH Aachen University, Germany,
barej@igm.rwth-aachen.de
2. Susanne Hoffmann,
scientific assistant at the
Chair of Structures and Structural Design,
RWTH Aachen University, Germany,
hoffmann@trako.arch.rwth-aachen.de
3. Christian Knief
former student assistant at the
Department of Mechanism Theory
and Dynamics of Machines,
RWTH Aachen University, Germany
4. Martin Trautz, professor, Ph.D.,
Head of the Chair of
Structures and Structural Design,
RWTH Aachen University, Germany,
trautz@trako.arch.rwth-aachen.de
5. Burkhard Corves, professor, Ph.D.,
Head of the Department of Mechanism Theory
and Dynamics of Machines,
RWTH Aachen University, Germany,
corves@igm.rwth-aachen.de

DEVELOPING SPATIAL ABILITY FOR QUALITY ENGINEERING EDUCATION

Maja ILIĆ¹ and Milena STAVRIĆ²

¹University of Banja Luka, BiH ²Graz University of Technology, Austria

ABSTRACT: Spatial intelligence, a relatively new branch of psychology, only began to be more extensively explored in the 1970's, and its role and relevance in assessing human intelligence is still not fully understood. The results of a survey completed last year (Kell et al 2013) suggest a significant impact of this kind of intelligence on creativity and success in patients with high scores in general intelligence tests. However, how relevant can the level of spatial ability be when it comes to specific professions that require the same degree of development of both analytical and divergent thinking?

Architecture is probably unique when it comes to this, since it necessitates competence in convergent thinking while insisting on a divergent approach to problem solving. Careful selection of candidates for the study of architecture will likely have a huge impact on their success during the study. The competencies required of prospective students are mostly based on the measurement of their ability and knowledge of mathematics, freehand drawing and spatial reasoning. Today's labour market requires much more of contemporary engineers. The recommendations of ENHSA – European Network of Heads of Schools of Architecture (Spiridonidis 2006) – are aimed at defining the general competence of architects that relates more to the ability of synthesis of acquired knowledge rather than its acquisition ("learning to learn" ability, capacity to apply a spirit of synthesis of ideas and forms, personal and social skills in expression and communication through speaking, writing and sketching, etc.). For that purpose, Jami Shah (Shah 2005) defines the term the *Designer Profile*, which includes a while range of skills needed to create a successful engineer (Imaginative Thinking (IT), Lateral Thinking (LT), Problem Definition Skill (PD), Visual Thinking (VT), Qualitative Reasoning/Abstract Vertical Thinking (AVT)). Thus, it seems that creative thinking is just as important for a candidate's success as their ability of analytical reasoning.

In addition to mathematics and freehand drawing tests, the spatial ability tests that have so far been used in university entrance tests generally focused on convergent thinking, with a single correct solution to the given task. To what extent are such tests successful in assessing the likelihood that a prospective student will excel in the study of architecture, which requires a comprehensive, creative and multilevel approach to problem solving?

This paper offers a comparative analysis of the tests of spatial ability used in the architecture entrance test at the Faculty of Architecture and Civil Engineering in Banja Luka in the last three years and of the matriculated students' achievement in a particular group of subjects during the first four years of study. Attention will be given to their overall academic achievement, along with that concerning in a group of subjects related to analytical thinking, and that concerning a group of design subjects requiring an integrated and creative approach.

The results will show the relation between spatial skills and success in specific study areas of the educational profile of architects, and provide recommendations for upgrading the methods of improving spatial ability in order to provide more adequate competences of candidates.

Further research should indicate the need to reassess the role of spatial ability at the university level, its role in the process of creative thinking, and should provide guidance for defining didactic meth-

ods for increasing the cooperation between convergent and divergent thinking by improving the level of spatial abilities. The aim will also be to point out the benefits of developing spatial ability at lower levels of education, and to examine the appropriate modules for implementation of study results in primary and secondary schools.

Keywords: spatial ability, creativity, entrance tests, divergent thinking, designer profile

1. INTRODUCTION

The development of technology and digital media in the last 30 years has significantly influenced the way that students learn at all educational levels. Worldwide there are numerous research studies that are dealing with teaching methods and development of new ways to acquire and use knowledge. The objectives of those researches are to improve teaching methods that would accommodate the needs of modern society more adequately. Recommendations of the European Union relating to the Educational policies and initiatives of the European Union highly value skills of creative thinking and multi-perspective approach in education. In line with these recommendations, the European Association of Architectural Education EAAE provides guidelines *New Directives - New Directions, Academic Direction of Our Schools in New Context* [1]. In this context, the importance of reviewing competence of architectural profiles with respect to the contemporary needs of the profession and the role of creative and critical thinking in the development of future engineers are highlighted.

The modern labor market, in addition to specific professional competence of engineers also requires an innovative approach to problem solving. Defining the *Designer Profile*, Shah states that convergent thinking on which is based the current education system is insufficient for the successful design: "... The science based regimen promotes convergent thinking and deductive reasoning through closed-ended problem solving. While these skills are extremely valuable to any engineer, they are insufficient for design. Design requires abductive reasoning, not only deductive; It re-

quires divergent thinking, not only convergent thinking; It requires creative thinking, not only critical thinking ... " [2] For conceptual design, Shah have identified four distinct skills: Lateral Thinking, Imaginative Thinking, Visual Thinking, and Qualitative Reasoning (Abstract Vertical Thinking). [2]

Among many competencies, at the beginning of the study of technical sciences, spatial ability of student is extremely important. This is important because these competencies provide a basis for the adoption of knowledge in specialized areas in further studies. Students who entered the study of architecture in Banja Luka often have difficulties in perception of space. Primary and secondary education does not pay enough attention to development of spatial skills, and therefore, leaves pretty narrowed space for their development at university level.

At the same time, universities are increasingly encouraged to provide greater opportunities for the development of creativity in the engineering profession [3]. Engineering profession demands recognition, verification and independent problem solving through teamwork. More importantly, it requires a demonstration of originality and critical thinking, creativity and innovation in developing new methodologies [4] [5]. Unfortunately, little has been done on integrating creativity in the teaching process in the adequate way.

Recent psychological research on the effects of spatial abilities has shown that spatial ability is very important for creativity. In the late 1970s, 563 intellectually talented 13-year-old students in the U.S. (according to SAT results) were tested for spatial ability. Thirty years later, they conducted a study on the importance of

the influence of spatial ability compared with mathematical and verbal abilities. The aim of this study was to test the hypothesis that spatial abilities play a unique role in the development of creative products [6]. Criterion of creativity was defined by the number of patents of respondents.

Creativity is a relatively new field of study, and there is still no consensus on what it actually is. Gilford spurred a serious notion of creativity as a scientific discipline [7] when he presented his views on the nature of creativity and its factors in 1950. He placed creativity in the framework of the intellect, and he explained the nature of creativity with factors of divergent production, later frequently referred to as divergent thinking. [8]

Divergent thinking is the process of generating different ideas or solutions related to a given problem. Divergent thinking is spontaneous and nonlinear process. In the other hand, convergent thinking is the ability to apply previously defined rules in order to find one correct solution to a given task. This process is systematic and linear [9].

Divergent thinking has become important in creativity research, due to its link with the production market and the need to continually generate new ideas and products. The creative industries need both types of thinking - divergent to reach a lot of different ideas and convergent to narrow that number down to those realistic and achievable in the market. The creative problem solving, or any complex problem solving, require the ability of permanent and parallel mixing of both ways of thinking. [8]

2. ASSESSING THE ABILITY OF FUTURE ENGINEERS

The first step in the education of future engineers is the selection of candidates for the study. In addition to mathematics and freehand drawing tests, the spatial ability tests that have been used so far in university entrance tests at Faculty of Architecture and Civil Engineering in Banja Luka generally focused on convergent thinking, with a single correct solution to the

given task. To what extent are such tests successful in assessing the likelihood that a prospective student will excel in the study of architecture, which requires a comprehensive, creative and multilevel approach to problem solving?

The task of the entrance exam in free drawing comprises of drawing a space composition comprising of solids. Relation to spatial composition, composing drawing on a sheet of paper, knowledge of perspective, proportion, ability to see the various basic geometric shapes, light source and legality of the existing lighting conditions, changes caused by different lighting conditions are measured in this way (Figure 1.)

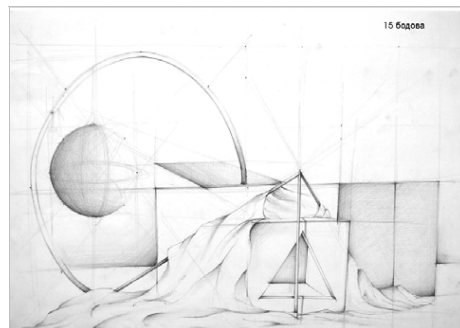


Figure 1. Example task at the entrance exam in freehand drawing

Entrance test in mathematics performs knowledge of mathematics acquired in high school. The test consists of five tasks that encompass an area of polynomials, equations, trigonometry, planimetry, and stereometry. In most cases, the mathematical problem is already defined (task 2-5), and for some tasks the candidate should define the problem (task 1 and 5).

Task 1.

The translators' team has 30 interpreters. Among them, 13 speak German, 12 speak French and 17 speak English. Five interpreters speak both German and English, 5 interpreters speak both German and French, and 5 interpreters speak both French and English. How many translators speak three languages?

Task 2.

Solve the inequality

$$\frac{3}{x^2-1} < -1.$$

Task 3.

Solve equation

$$3^{x+1} + 18 \cdot 3^{-x} = 29.$$

Task 4.

Show that the expression

$$\operatorname{tg} 15^\circ + \operatorname{ctg} 15^\circ$$

results with round number and which number is that

Task 5.

Determine the point on the y-axis at which a tangents drawn to the parabola $y = x^2$ are perpendicular to each other.

Figure 2. Example of entrance test in mathematics

Current entrance tests of spatial ability at Faculty of Architecture and Civil Engineering in Banja Luka focuses on solving spatial tasks by measuring mostly factors of spatial visualization, mental rotation and mental cutting. Unlike the task used in standardized tests of spatial ability, these tasks are at a higher level and are partly designed to have multiple choice (Figure 3), and in part require the candidate to generate their own solutions. However, there is still one correct answer (Figure 4).

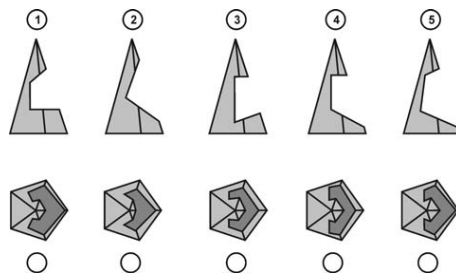


Figure 3: Example of entrance examination task in descriptive geometry and perspective. Text of the task: *The drawing shows five pyramids from which some parts have been cut out. At the bottom of the drawing are given top views of pyramids. In the circles write the number of the pyramid that suits a given view from above.*

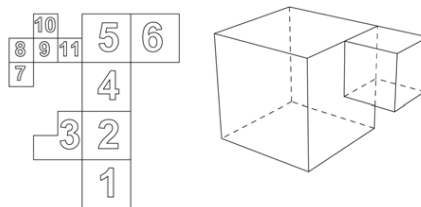


Figure 4: Example of entrance examination task in descriptive geometry and perspective. Text of the task: *On the left drawing shows developed surface of the object from the right. Each surface of the object is numbered 1-11. Taking into consideration the visible lines of the object and orientation of the numbers, draw the adequate numbers on the object given on the right. Numbers are drawn on the outer surface of the object.*

The aim of this research was to determine the correlation of each of these tests individually with students' academic success of the first two years of study. It was noted that in all tests dominated convergent thinking, where there is one correct answer. Special attention will be brought to the spatial ability tests in order to

determine its correlation with success in specific areas of the curriculum.

In addition, there have been organized preparatory classes for entrance exam. Preparatory classes are fifteen-day long intensive course in all three areas of testing and during that time, students are working on entrance exams held in previous years.

Research conducted at the Faculty of Architecture and Civil Engineering in Banja Luka on the influence of preparatory course and the results of the entrance examination showed a high correlation when it comes to math and freehand drawing, while in spatial ability there was not significant influences [10]. Based on the results the authors concluded that the tasks of spatial ability is designed to test the ability and inclination of candidates, and that knowledge of certain areas that candidates gain in the high school does not necessary bring advantage. This speaks to the fact that secondary education is not given enough attention to the development of spatial ability.

3. RESEARCH

The research of this paper was conducted on a sample of 181 students of architecture through three generations of registered candidates. After two years of study the evaluation of the students during those two years were analyzed and the results were compared with the results achieved in the entrance exam. To determine the correlation of entrance examination with success during the study, subjects to which the analysis was performed were divided into three groups:

- A. Subjects in which dominates mathematical and logical thinking
- B. Subjects in which dominates the spatial-visual thinking
- C. Subjects in which an integrated approach to thinking is required (designing and constructing)

Due to the fact that success of candidates in individual cases depends on a number of factors (the subjectivity of teachers, teaching

methods, program objectives, the structure and method of evaluation...) data were taken through three successive generations of students (2009-2011). Possibility to express creativity depends on the teachers and curriculum, but it is possible to group the subject in terms of dominating way of thinking. Subjects in which dominates mathematical and logical thinking include mathematics, physics, statics and mechanics where the problem is pre-defined and students are required to focus on the most common one possible solution to the problem. In the group of subjects in which dominates the visual-spatial thinking are architectural graphics, painting, visualization and modeling. Subjects in whom an integrated approach is required are constructing and architectural design, where the task is not strictly defined, and it requires the production of a large number of possible solutions.

Method of regression was used in statistical data analysis using SPSS software.

4. RESULTS

Taking into account the results of all three segments of entrance exam (Psc, Pmm, Png), entrance exam has significant predictive value for a group of mathematical and logical subjects ($AvML = .46$). Entrance exam from mathematics (Pmm) provides a significant level of 0.01, while the entrance exam in freehand drawing (Psc) at the level of 0.05.

Taking into account the results of all three segments of entrance exam (Psc, Pmm, Png), entrance exam has significant predictive value for a group of visual spatial subjects ($AvPV = .52$). Entrance exams in mathematics and freehand drawing (Pmm, Psc) provide a significant level of 0.01, while the entrance exam in spatial ability (Png) at the level of 0.05.

Taking into account the results of all three segments of entrance exam (Psc, Pmm, Png), entrance exam has significant predictive value for the group of construction design subjects ($AVKO = .45$). Entrance exams in mathematics and freehand drawing (Pmm, Ps) provide a significant level of 0.01, while the entrance exam

in spatial ability has no significance.

Taking into account the results of all three segments of entrance exam (Psc, Pmm, Png), entrance exam has significant predictive value for the group of architectural design subjects (AVP = .45). Entrance exams in mathematics and freehand drawing (Pmm, Psc) provide a significant level of 0.01, while the entrance examination in spatial ability has no significance.

Taking into account the results of all three segments of entrance exam (Psc, Pmm, Png), entrance exam has significant predictive value for overall academic achievement of students (AvTotal = .51). Entrance exams in mathematics and freehand drawing (Pmm, Psc) provides a significant level of 0.01, while the entrance examination in spatial ability has no significance.

All variables correlate with entrance exams and they do not interfere with each other, which means that they measure different abilities.

5. DISCUSSION

The results showed that the entrance exam has significant predictive value for all groups of subjects. The entrance exams in mathematics and freehand drawing have significant predictive value, while the entrance examination of spatial ability has significance only in the mathematical-logical group of subjects.

This speaks of the necessity of inclusion of divergent of thinking in solving spatial problems. The inclusion of creative thinking in the process of solving spatial problems contributes to the development of strategies for solving spatial problems. The main contribution would be in producing a number of possible solutions that could contribute in developing divergent thinking with students. A proposed example of such tasks is shown in Figures 5. and 6.

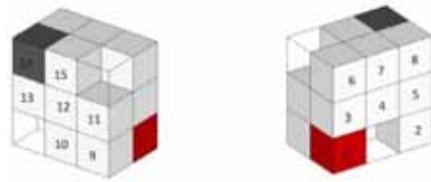


Figure 5. Text of the task: Write down the steps of moving red cube (1) in a position of gray cube (14) in a shortest way. Cubes are moving within the rectangular wire cuboids. Answer: There are several possible solutions.

The minimum number of moves is 13
(1-3-4-5-2-1-1-2-13-1-15-14-1), authors M. Ilić
and M. Stavrić

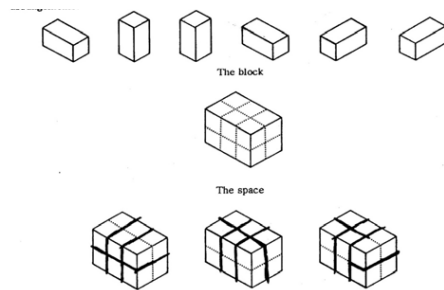


Figure 6. Text of the task: There are 6 identical blocks. Draw as many possible solutions to how these blocks may be arranged to give the cuboids. Draw visible and invisible lines.

[11]

6. CONCLUSION

Tests of freehand drawing, math and spatial abilities are part of the entrance examination at the Faculty of Architecture and Civil Engineering in Banja Luka. Analysis of the results of these examinations for the past three years and academic success of students during first two years of studies showed a statistically significant correlation of results in mathematics and free drawing with success during the study, while spatial ability tests showed no statistically significant correlation. Many factors can influence the lack of correlation. One of the rea-

sons can be in preparatory course held prior the examination test where students acquire a routine exercise in solving certain types of problems, while their spatial abilities remain undeveloped. It is necessary to develop spatial abilities at lower educational levels, because at the time when students enroll in college they are already in deficit. Entrance examination not necessary need to be eliminating factor in enrollment, but it can be a detector of potentials and weaknesses in students on which they should work in further education. These options certainly depend on the economic and staff situation at the faculty, because with individual students it is necessary to pay more attention to reduce the initial deficits.

Tests of spatial ability can be improved to better suit the needs of modern engineers in a way to define tasks with several possible solutions in order to develop divergent thinking in students. This principle can be integrated into other curriculum segments in college in order to long-term process of thinking development.

ACKNOWLEDGMENTS

Statistical data were analyzed with help of Milica Drobac, psychologist at the Philosophical Faculty in Banja Luka.

REFERENCES

- [1] New Directive - New Directions...What is the Academic Direction of our Schools in this New Context?, in *9th Meeting of Heads of European Schools of Architecture*, Crete, 2006.
- [2] J. Shah. Identification, measurement, and development of design skills in engineering education. In *Proceedings of ICED 05, the 15th International Conference on Engineering Design*, Melbourne, 2005.
- [3] C. Baillie and P. Walker. Fostering creative thinking in student engineers. *European Journal of Engineering Education*, 23(1): 35-44, 1998.
- [4] M. Shaw. *Engineering Problem Solving: A Classical Perspective*. Norwich, NY: Noyes Publications, 2002.
- [5] Z. Liu i D. Schönwetter. Teaching Creativity in Engineering Education. *International Journal of Engineering Education*, 20(5): 801-808, 2004.
- [6] D. Lubinski, C. Benbow, R. M. Webb and A. Bleske-Rechek. Tracking Exceptional Human Capital Over Two Decades. *Psychological science*, vol. 17, no. 3, pp. 194-199, 2006.
- [7] J. Guilford. Creativity. *American Psychologist*, tom 5, pp. 444-454, 1950.
- [8] M. A. Runco. Reasoning and Personal Creativity. In *Creativity and Reason in Cognitive Development*. New York, Cambridge University Press, 2006, pp. 99-116.
- [9] J. P. Guilford. The nature of human intelligence, New York: McGraw-Hill, 1971.
- [10] LJ. Preradović, M. Dodig, M. Vinčić, B. Antunović, S. Kosić-Jeremić and D. Borojević. Uticaj srednjoškolskog obrazovanja i pripremne nastave na uspjeh kandidata pri polaganju kvalifikacionog ispita na Arhitektonsko-građevinskom fakultetu Univerziteta u Banjaluci tokom 2012. godine. Arhitektonsko-građevinski fakultet Univerziteta u Banjoj Luci, Banja Luka, 2013.
- [11] O. Akin, Spatial Reasoning of Architecture Students with Simple Three-Dimensional Arrangements. *Journal of Design Research*, 9(4): 339-359, 2011.
- [12] S. A. Sorby. Developing 3-D Spatial Visualization Skills. *Engineering Design Graphics Journal*, 63(2): 21-32, 1999.
- [13] H. J. Kell, D. Lubinski, C. P. Benbow and J. H. Steiger. Creativity and Technical Innovation: Spatial Ability's Unique Role. *Psychological Science*, 24(9): 1831-1836, 2013.

ABOUT THE AUTHORS

1. Maja Ilić is a senior teaching assistant at the Faculty of Architecture and Civil Engineering in Banja Luka (BiH) on subjects of Descriptive Geometry, Architectural Graphics and Visualization and Modeling. Her research interests are in assessing and training for spatial skills and and computer aided design.

2. M. Stavrić is an Ass. Prof. at the Faculty of Architecture, Institute for Architecture and Media - Graz University of Technology. Her research topic is the in complex geometry in architecture.

THE DEVELOPMENTS OF MODELING TOOLS FOR PIN-BY-PIN PRECISE REACTOR SIMULATIONS

Yan MA ¹, Yuanguang FU ², Shu LI ¹, Gang LI ¹ and Li DENG ¹

¹ The Institute of Applied Physics and Computational Mathematics, China

² China Academy of Engineering Physics Software Center of High Performance Numerical Simulation, China

ABSTRACT: In order to develop large-scale transport simulation and calculating method (such as simulation of whole reactor core pin-by-pin problem), the Institute of Applied Physics and Computational Mathematics (IAPCM) developed the neutron photon coupled transport code JMCT and the toolkit JCOGIN. To create physical calculation model easily and efficiently can essentially reduce problem solving time cost. Currently, lots of visual modeling programs have been developed based on different CAD systems. In this article, the developing idea of a visual modeling tool based on field oriented development was introduced; Considering the feature of physical modeling, fast and convenient operation modules were developed; In order to solve the storage and conversion problems of large-scale models, the data structure and conversion algorithm based on the hierarchical geometry tree were designed; The automatic conversion and generation of physical model input file for JMCT were realized. By using this modeling tool, the DAYAWAN whole-core physical model was created, and the transformed file was delivered to JMCT for transport calculation. The results were compared and validate the correctness of the visual modeling tool.

Keywords: Field Oriented Modeling, Pin-by-Pin Simulation, Hierarchical Geometry Tree, GDML.

1. INTRODUCTION

The Monte Carlo (MC) transport calculation method is widely used in nuclear design and assessment, especially in neutronics analysis such as reactor core design, reactor shielding design and reactor safety analysis. It simulates large amount of neutrons' transport histories in different mediums. It is necessary to give a geometry description of mediums. With the numerical and engineering simulation progressing, results with high confidence can be achieved by using high performance scientific and engineering computing applications. In order to achieve results with high precision by MC calculation, much more detailed geometry of model should be given. Take the DAYAWAN reactor core as an example, the core consists of 157 fuel assemblies, each assembly has 17X17 fuel pins, and in order to get a precise

power distribution of the core, each pin should be divided into 16 sections. Then the total solid number of this model reaches 725968.

In order to simulate reactor pin-by-pin model such as DAYAWAN, the Institute of Applied Physics and Computational Mathematics (IAPCM) developed a neutron-photon coupled MC code JMCT (J Monte Carlo Transport), based on the CERN's MC program Geant4. Although Geant4 supply a complete geometric description capacity, it lacks a geometric modeling tool. Users should build models by manually writing text files. A simple model can be easily built by this method, but many problems will arise when dealing models with complex geometries: 1) Modeling efficiency is low and building complex models costs too much time. 2) Complex Boolean operations are hard to realize. 3) Error checking is hard in complex models. 4) Modifying subtle error is trivial and

fallible. As MC algorithm develops, the manual modeling method becomes a restriction. In order to handle models with complex geometry such as the DAYAWAN pin-by-pin reactor core, we developed a modeling tool JLAMT (J Large-scale Automatic Modeling Tool).

2. THE VISUAL MODELING TOOL JLAMT BASED ON FIELD ORIENTED DEVELOPMENT

In order to improve 3D modeling efficiency, it is necessary to develop a visualized modeling tool based on CAD kernel. UG is one of the most advanced CAD, CAM software. Therefore we did the secondary development on UG which forms JLAMT. JLAMT will automatically export GDML (Geometry Description Markup Language) type files which can be used by JMCT and other MC codes.

There are great differences between the CAD model and the physical model used by MC code. The data structure to store geometry is completely different. What's more, although by using CAD we can build detailed geometry of models, they may not satisfy MC calculation. Not only the geometry, but only the materials and other neutronic parameters are also needed. Also, MC codes do not allow overlap between solids. Therefore, it is necessary to do conversions between these types. Lots of MC modeling programs are designed and developed mainly focusing on the conversion algorithm, but special modeling needs are neglected.

It is a fact that although a reactor core is complex, it has a lot of repeated structures. Based on this field special knowledge and needs, we developed modules aiming at fast reactor core modeling on JLAMT. It uses parameterized methods to simplify the reactor assembly modeling procedure. The assembly modeling UI is shown in Figure 1.

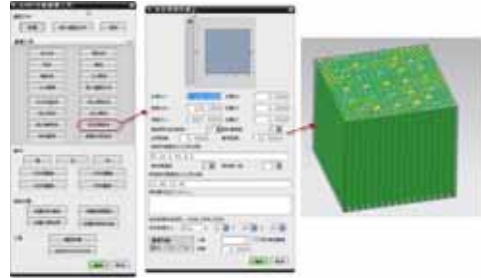


Figure 1: The reactor assembly modeling tool of JLAMT.

We developed JLAMT mainly focusing on MC code users' special needs and habits of geometry and physical modeling. The program supplies a guided interface, which makes the user to set parameters (such as solid type, location, direction, name, material, neutronics parameters and so on) step by step. The user may build MC models without trainings by using JLAMT.

3. GEOMETRY DESCRIPTION METHOD OF CAD AND JMCT

There are two main methods to store 3D models, the CSG and the Brep. The basic idea of the CSG method is: any arbitrary 3D models can be achieved by using series of complex Boolean operations between simple voxel. The CSG method can be regarded as an ordered binary tree. The leaf nodes are basic voxels, the intermediate nodes are Boolean solids, and the root node is 3D model. The main advantage of CSG method is that its data structure is simple, the memory consume is small, and the memory management is easy. Many MC codes choose CSG as their geometry description method, including Geant4 and JMCT.

In CSG method, a physical model is often formed by different types and size of solids. To define a complete physical model, the types of solids, the relative position and location between them, and the materials and neutronics parameters should be given completely. A complete description of a physical in Geant4

should use three layers definition:

1. Solids. Define the type and size of objects.
2. Logical Volume. Define the materials of solids.
3. Physical Volume. Define the positions and rotations of logical volumes

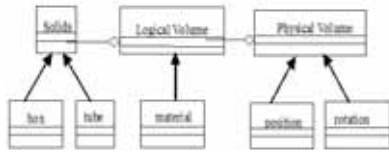


Figure 2: Three layers definition in Geant4.

The three layers definition makes Geant4 a flexible ability in geometry description. Solid layer defines the basic geometry information. Solids with same shape but different materials can be repeated by using Logical Volume layer. Logical Volumes with same geometry and material but different placements can be repeated by using Physical Volume layer. By using three layers definition, it's much easy to describe repeated structure, and it's much efficient and clear to build and manage large-scale models.

In CSG method, different types of solid determine the geometry description ability. Geant4 supplies not only some basic voxels (such as box, sphere, cylinder, cone, etc.), but also some special voxels (such as ellipsoid, torus, twisted box, etc.). Based on these voxels, complex Boolean operation can be done. These guarantee Geant4's ability of building complex geometry. All types of voxels and Boolean operations are supported by JMCT.

Another method Geant4 use to describe geometry is the mother-daughter relationship between solids. This property mainly describes the inclusion relations between two volumes. Fig.3 is an example: volume "a" contains volume "b", "c", "d", then volume "a" is the mother volume and volume "b", "c", "d" are the daughter volumes. The daughter volume such as "c" can also have its daughter. Though the

number of hierarchy levels is not limited, there must be only one root volume "world". These volumes form a hierarchy geometry tree. This method greatly simplifies the description of repeated structure. Take Fig.3 as an example, if one needs to describe another volume "a1", which has same geometry with "a", he should describe branch "world-a-c-e" in detail, and then repeat "a" once as "a1", he needs not to describe "world-a1-c1-e1" branch in detail because volumes inside "a1" (such as "c1", "e1", etc.) are implicitly repeated. In DAYAWAN reactor core model, geometry is same among assemblies, by using mother-daughter relationship, repeated structure modeling work can be greatly reduced. Also, the MC transport calculation does not need to simulate in global geometry, but to simulate in geometry branch. Computing time can be greatly reduced.

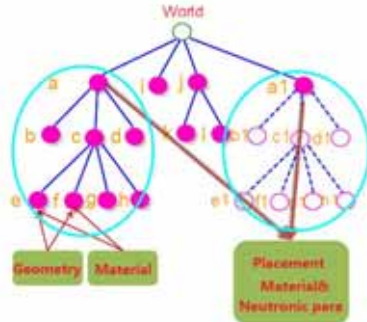


Figure 3: Sketch map of hierarchy geometry tree.

Generally, most of CAD kernels use Brep method to describe and store solids. This method is mature and has no ambiguity. Based on solids, surface, loop, edge, point, the Brep method completely describes the information of all elements and topological relationships between them, which makes it capable of describing any arbitrary shape. Moreover, commercial CAD software often uses their own file format or industry standard format to store and exchange CAD data. Therefore, there are big differences

in geometry description between CAD and JMCT. JLAMT works as a pre-processing unit for JMCT. It converts CAD data in UG to CSG data based on GDML, which can be read by JMCT. In order to arrange data properly, we developed our data structures and algorithms to finish conversion job. JLAMT can now handle several voxels such as box, sphere, cylinder, cone, etc. Any arbitrary hierarchy trees can be set by users. Any placement and rotation as well as complex Boolean operations can be converted. Large-scale models with amounts of placement data and complex mother-daughter relationship such as DAYAWAN core model can be converted to GDML correctly.

One would face a series of problems if he builds a large-scale complex model on PC. The memory may be insufficient, and the interactive or display speed may be unbearable. Therefore, we developed a new method, “implicit modeling” based on hierarchy geometry tree description. Take the DAYAWAN reactor core as an example, shown in Fig.4, the core consists of 157 fuel assemblies, and each assembly has 17X17 fuel pins, and each pin has three layers; Each assembly is divided in 16 assembly block axially. So in the core there are 2512 assembly blocks, and each block contains the same 17x17 pin structure. The total solids number is 2177904, which is too large to model and display. If we use implicit modeling, we only need to build 17x17 pin structure in one block in detail, for the other 2511 blocks, we only build simple boxes without detailed geometry in them. In JLAMT, we label these 2511 repeated blocks and make them reference to the detailed one. In the final converted GDML, 17x17 pin structure in each block is implicitly repeated, while in the CAD model only one block’s detailed geometry is displayed. The implicit modeling can be explained by Fig.3, volume “world” can be regarded as core, volume “a” is one assembly block, volume “b”, “c”, “d” are fuel pins, and volume “e”, “f” are layers in each pin. This branch is a geometry subtree which contains complete 17x17 pin structure geome-

try information. The other 2511 blocks are volume “a1”, “a2”... which have same geometry as volume “a”. They can be built by only repeating the root node of a subtree, and the geometry information inside is implicitly repeated. By using this method in JLAMT, we guarantee that models with millions of repeated solids can be displayed by thousands of solids in CAD, and detailed geometry information can be converted to GDML.

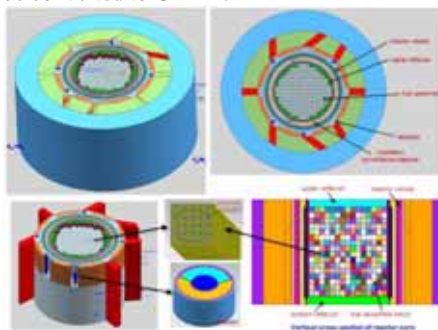


Figure 4: The rendering picture of DAYAWAN reactor full-core.

4. GEOMETRY DESCRIPTION FILE FORMAT GDML

JLAMT export GDML file format as geometry description to JMCT. GDML is developed based on XML, and uses CSG method. Based on syntax of XML, GDML gives its geometry definition rules which include how to define types, sizes, placements, Boolean operations, mother-daughter relationships, and materials of solids.

A GDML file consists of five parts:

1. `<define>...</define>`: It defines constants, placement parameters (location, rotation). This part can be referenced by latter part of the file.
2. `<material>...</materials>`: It defines the kind of mixture filled in solids.
3. `<solids>...</solids>`: It defines solids types, including 24 kinds of basic voxel and their sizes. It also defines solids

built by complex Boolean operations (intersection, union, subtraction).

4. `<structure>...</structure>`: It defines logical volumes and physical volumes, and how hierarchy geometry is organized by these volumes. It also defines the materials of logical volumes, and the placements of physical volumes.
5. `<setup>...</setup>`: It defines world, the root of all geometry subtrees.

These five parts are relatively independent. Data in different parts can refer to each other. The same constants, materials parameters, and hierarchy trees need only to be described once, and they can be referenced many times as repeating. GDML files are suitable to describe large-scale model.

JLAMT is developed on UG. Users can build models by using UG's interaction functions. JLAMT can automatically convert CAD models to GDML files for users to do further neutron transport calculations.

5. CONCLUSIONS

We introduce the modeling tool JLAMT specially designed for MC transport geometry modeling. It is developed based on feature modeling technology of CAD. The field based development tool JLAMT supplies faster and more convenient operations than manually modeling, which shortens the user's learning cycle. Based on UG, JLAMT supplies a friendly user interface, which makes the user easy to build, edit, and modify geometry. It also supplies multiple ways for users to set materials and other neutronic parameters, which makes the user efficient to build physical models. JLAMT can automatically convert large-scale models with complex geometry in CAD to GDML file, the hierarchy is used as a natural way and as an efficient way of organizing the geometry, which can be used by CSG based MC neutron transport program to do further calculation. The results are good; this also means the auto-modeling tool is a feasible and effective way for JMCT model input.

ACKNOWLEDGMENTS

This work was performed under the National High Technology Research and Development Program of China (No. 2012AA01A303) and the CAEP fund (No. 2011A0103006, 2012A0102005).

REFERENCES

- [1] DENG Li, LI Gang, ZHANG BaoYin, et al. SIMULATION OF THE FULL-CORE PIN-MODEL BY JMCT MONTE CARLO NEUTRON-PHOTON TRANSPORT CODE, p427-p433, M&C2013, Sun Valley, Idaho, May 5-9, 2013.
- [2] ZHANG BaoYin, LI Gang, DENG Li, JCOGIN: a combinatorial geometry Monte Carlo particle transport infrastructure, High Power Laser and Particle Beams, 2013, 25(1), p173-176, MianYang, CHINA.
- [3] WU YiCan, LI Ying, LU Lei, et al. Research and Development of the Automatic Modeling System for Monte Carlo Particle Transport Simulation, Chinese Journal of Nuclear Science and Engineering, 2006. 26(1), p20-27, Beijing, CHINA
- [4] S. Agostinelli, J. Allison, et al. Geant4--a simulation toolkit. Nuclear Instruments and Methods in Physics Research, NIM A 506 (2003) 250-C303. See GEANT4 Web page: <http://cern.ch/geant4>.
- [5] J. Apostolakis, R. Chytrcek, G. Cosmo, et al. Recent Developments and Upgrades to the Geant4 Geometry Modeler. Proceedings of the CHEP 2006 Conference, See also the GEANT4 web page.
- [6] Geant4 Collaboration, The GEANT4 User Guide for Application Developers[R]. Accessible from the GEANT4 web page.
- [7] Geant4 Collaboration, GDML USER'S GUIDE[R]. See also the GDML web page: <http://cern.ch/gdml>.

ABOUT THE AUTHORS

1. Yan MA, is an associate professor at the Institute of Applied Physics and Computational Mathematics of Beijing. Her research interests include computer graphics and visualization in scientific computing. Her e-mail and postal address is as follows: mayan2301@gmail.com, No.2, Fenghao East Road, Hai-dian District, Beijing 100094, China

2. Yuanguang FU, is an assistant research at Software Center of China Academy of Engineering Physics. His research interests include nuclear reactor physics and modeling. His e-mail and postal address is as follows: fyg1989@126.com, No.6 Huayuan Road, Hai-dian District, Beijing 100088, China

3. Shu LI, is an associate professor at the Institute of Applied Physics and Computational Mathematics of Beijing. His research interests include nuclear reactor physics and Monte Carlo method. His e-mail and postal address is as follows: li_shu@iapcm.ac.cn, No.2, Fenghao East Road, Hai-dian District, Beijing 100094, China

4. Gang LI, is an associate professor at the Institute of Applied Physics and Computational Mathematics of Beijing. His research interests include nuclear physics and reactor neutronics. His e-mail and postal address is as follows: li_gang@iapcm.ac.cn, No.2, Fenghao East Road, Hai-dian District, Beijing 100094, China

5. Li DENG, is a professor at the Institute of Applied Physics and Computational Mathematics of Beijing. His research interests include method and application software for MC simulation. His e-mail and postal address is as follows: deng_li@iapcm.ac.cn, No.2, Fenghao East Road, Hai-dian District, Beijing 100094, China

DYNAMIC GEOMETRY AND WEBSITE SETUP BY AUTOMATIC OBJECT RECOGNITION OF FREE-HAND DRAWINGS AND SCANS

Karl-Heinz BRAKHAGE
RWTH Aachen University, Germany

ABSTRACT: Computer software, websites, web applications and the usability on tablets and smart-phones have the potential to make significant improvements in how geometry is learnt and taught. As with any approach to teaching, the educational use of information and communication technology as well as computer software needs to be well thought through and carefully planned. Sometimes it is useful to have tools that automatically recognize given curves as lines, circles or conics and splines. In this paper we will concentrate on automatic object recognition from free-hand drawings and / or scans and the conversion into scalable vector graphics. These can be used for animations or exported as vector graphics in various formats. We will give an overview of the analytical background and implementation issues. Furthermore we will demonstrate its use on a couple of selected examples from various areas.

Keywords: Teaching, Descriptive Geometry, Dynamic Geometry, Splines, Approximation, Fairing, Numerical Analysis

1. INTRODUCTION

Nowadays the abilities of computers and software have a significant influence on teaching and learning geometry. Computer software, particular that known as Dynamic Geometry Software, websites, web applications and the usability on tablets and smartphones have the potential to make significant improvements in how geometry is learnt and taught. Our view is that teachers should now have at their disposal an appropriate variety of equipment from which to select. As with any approach to teaching, the educational use of information and communication technology as well as computer software needs to be well thought through and carefully planned. In this paper we will concentrate on some special topics of *WinCAG* (Windows version of Computer Aided Geometry) which has been especially designed for teaching and learning geometry (see [2]). With this software package both simple constructions and complicated dependencies can

easily be animated. Furthermore the drawings can be exported in well suited formats for further use in publications or on web pages. Here we will particular focus on some special capabilities for automatic object recognition from free-hand drawings and / or scans. Sometimes it is useful to have tools that automatically recognize given curves as lines, circles or conics and splines. We have implemented that in *WinCAG* in two ways. For external data we can import pixel oriented bitmaps or polylines in form of coordinates. Furthermore the system has a module that generates traverses by free-hand mouse-drawing.

To clearly demonstrate the goal of our efforts we give an illustrating example. Figure 1 shows a screenshot of such a scenario. We have three different traverses. These were *recognized* by the system as a line, a circle and an ellipse. Figure 2 shows the reconstruction of the corresponding scanned plot of figure 1. A zoom clearly shows a pixelated look in case of figure 1 and the *exact*

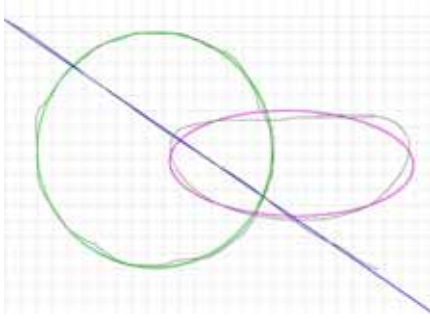


Figure 1: Screenshot of an object recognition: line, circle and ellipse (with raster).

curves in figure 2. Additionally hyperbolas and

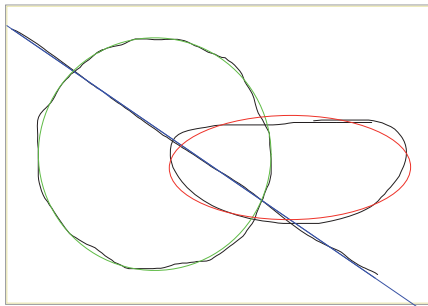


Figure 2: Reconstruction of figure 1 (without raster).

parabolas will be recognized. For this purpose the user has to give tolerances for each object. Since for instance an ellipse normally fits better than a circle, the tolerance for circles must be chosen larger than that for ellipses. The system stores the data traverses and the user can force it to compute a given kind of curve. The fall back option is a smooth B-spline. The curves corresponding to the input in form of polylines or point clouds are computed by least squares solutions. For conics the fit by a linear implicit form does not yield the optimal (non-linear) least

squares solution given by the parameterized form. Thus for an initial guess from a linearized scaled problem we determine the parameters for the data points and make some Gauss-Newton steps to improve our result. If we are *close to* a parabola the system computes a least squares solution for this case, too. Then the user can optionally decide to take the parabola. If the distances to all of the above objects are too large the system will compute a B-spline approximation. Figure 3 shows an example for that situation. We have plotted the control polygon for the B-spline (of order 4), too. Additionally we have marked the data points because the data polygon nearly coincides with the spline.

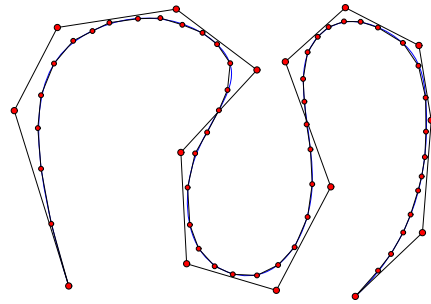


Figure 3: PDF export of an object recognition: B-spline.

The geometric objects of the above processes can either be used within the software package **WinCAG** for animations or exported as vector graphics in various formats, for instance Encapsulated Postscript (EPS), PDF and Scalable Vector Graphics (SVG). EPS and PDF files are for use in printed publications. SVGs can be used for websites. The major advantage of all three formats is the very small size of the files and the possibility to zoom without grid pattern effects (see figure 1). This is achieved via highly accurate numerical approximation of the given geometrical entities using objects supported in the target format.

We have seen above that raster images like screen shots cannot be scaled up in size without the loose of sharpness – they will look pixelated. Producing files (in the background) with a sufficient resolution for high quality prints results usually in very large files. Thus formats based on raster images are useless for websites – at least if we need higher resolutions. On the other hand vector graphics formats can present lines, curves, fonts etc. at essentially arbitrary precision in different colors and sizes in *small* files. Thus the conversion into these formats yields a significant quality improvement for the representation of curves and other plots.

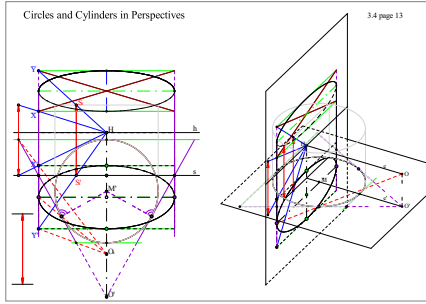


Figure 4: Advanced stage of an example for a perspective view with all layers on.

Figure 4 shows a more complex example. Here we have turned all layers on. Details will become clear by zooming into different areas of the plot.

The rest of this paper is organized as follows. First we give some basic notations and properties of Bézier and B-spline curves. Next we give some details on the used approximation strategies and algorithms for splines. A short summary and an evaluation regarding computation time and accuracy of the algorithms usually applied for these purposes will be given. Then we show how we force the system to favour special entities such as ellipses, hyperbolas or parabolas. Finally a we make some notes on the module that converts bitmaps to vector graphics.

2. SOME BASICS ON BÉZIERS AND B-SPLINES

Bézier curves of order n are given by $n + 1$ control or Bézier points \mathbf{p}_i $i = 0, \dots, n$ (here $m = 2$), that build the so-called control polygon, are given for $t \in [0, 1]$ by

$$\mathbf{x}(t) = \sum_{i=0}^n B_i^n(t) \mathbf{p}_i \quad (1)$$

with the Bernsteinpolynomials

$$B_i^n(t) = \binom{n}{i} t^i (1-t)^{n-i} \quad (2)$$

It is easy to verify

$$\begin{aligned} \mathbf{x}(0) &= \mathbf{p}_0 & \mathbf{x}(1) &= \mathbf{p}_n \\ \mathbf{x}'(0) &= n(\mathbf{p}_1 - \mathbf{p}_0) & \mathbf{x}'(1) &= n(\mathbf{p}_n - \mathbf{p}_{n-1}) \end{aligned} \quad (3)$$

Geometric continuity is an important property of the shape of curves and surfaces. For dynamic applications, e.g. path of a cutter tool for NC milling, we need curvature continuity. This is due to the physical effect of mass inertia. For our purposes here it is enough to claim tangent continuity. Thus we only define GC^1 -continuity in terms of the parametrization. Two curves $\mathbf{x}(t) = \mathbf{x}(t_0, t_1)$ and $\mathbf{y}(s) = \mathbf{y}(s_0, s_1)$ are said to be GC^1 -continuous at the junction $\mathbf{x}(t_1) = \mathbf{y}(s_0)$ (that is called GC^0 -continuity) if there exists a γ with

$$\mathbf{x}'(t_1) = \gamma \mathbf{y}'(s_0) \quad (4)$$

From (3) it is easy to see that the composition of two Bézier curves $\mathbf{x}^1(t)$ and $\mathbf{x}^2(t)$ is GC^1 -continuous, iff $\mathbf{p}_{n-1}^1, \mathbf{p}_n^1, \mathbf{p}_0^2$ and \mathbf{p}_1^2 lie in a straight line.

One main advantage for interpolation and approximation with curves like (1) is the linearity in the control points \mathbf{p}_i . B-spline curves have these property, too. They are given by

$$\mathbf{x}(t) = \sum_{i=0}^N N_{i,p,T}(t) \mathbf{p}_i \quad (5)$$

where $N_{i,p,T}(t)$ is the i -th normalized B-spline function of order p (degree $p - 1$) corresponding

to the generally non-uniform knot vector T $t_0 t_1 \dots t_{N_p}$. We usually assume that T is clamped, i.e., $t_0 = t_{p-1}$ and $t_N = t_{N-1}$. For the sake of simplicity we write N_{ip} instead of N_{ipT} since it becomes clear from the name of the function argument what the knot vector is. Surfaces are represented by B-spline tensor products of the form

$$\mathbf{x}(u, v) = \sum_{i=0}^N \sum_{j=0}^M N_{ip}(u) N_{jq}(v) \mathbf{p}_{ij} \quad (6)$$

and the Bézier representation is

$$\mathbf{x}(u, v) = \sum_{i=0}^N \sum_{j=0}^M B_i^n(u) B_j^m(v) \mathbf{p}_{ij} \quad (7)$$

These representation are linear in the control points, too. The same applies for stationary subdivision curves and surfaces. The algorithms presented below can be used for all these curve and surface classes.

Further information on Bézier- and B-splines can be found in standard literature on CAGD (Computer Aided Geometric Design), e.g. [5]. Furthermore several web page with applets regarding splines can be found. Here are some examples

- Wikipedia-Bezier: <http://de.wikipedia.org/wiki/B%C3%A9zierkurve>
- Prautsch-Applets: <http://i33www.ibds.uni-karlsruhe.de/applets/mocca/html/noplugin/inhalt.html>
- Farin-Applets: <http://www.vis.uni-stuttgart.de/~kraus/LiveGraphics3D/cagd/index.html>

3. APPROXIMATION OF CURVES AND SURFACES

Postscript, PDF and SVG have the entities quadratic and cubic Bézier curves but no higher

degrees. Therefore we will restrict ourselves to polynomial degree two and three in the following if it reduces technical investments for the description.

B-splines can be split into Béziers by successive knot insertion. If we use cubic B-splines we get GC^2 -continuity and in general approximations with smaller errors than those from GC^1 -continuous Bézier approximations. But the linear systems of equations we have to solve are larger.

To be more precise we have to introduce some technical notations. For an optimal approximation of a given curve $\mathbf{y}(v)$ $v \in [v_{min}, v_{max}]$ by a Bézier-spline we have to determine the control points \mathbf{p}_i (see (1)) in such a way that

$$\max_{v \in [v_{min}, v_{max}]} \min_{t \in [0, 1]} \|\mathbf{y}(v) - \mathbf{x}(t)\|^2 \quad (8)$$

is minimized. For B-splines the only difference is that in (8) now $t \in [t_{p-1}, t_N]$. Since these problems are too hard to solve we switch to discrete interpolation or approximation problems. For Béziers the whole interval $[v_{min}, v_{max}]$ is normally a-priori splitted into a couple of subintervals. For simplicity we use the interval $[v_l, v_r]$ for these subintervals and define

$$v : s \mapsto v_l + s(v_r - v_l) \quad (9)$$

Since we want to construct over all GC^1 -continuous curves we further assume that for all parts $\mathbf{y}(v_l)$, $\mathbf{y}(v_l)$, $\mathbf{y}(v_r)$ and $\mathbf{y}(v_r)$ are given or computable. Then for quadratic Béziers from (3) we get

$$\begin{pmatrix} \mathbf{p}_0 & \mathbf{y}(v_l) \\ \mathbf{p}_2 & \mathbf{y}(v_r) \end{pmatrix} \begin{pmatrix} \mathbf{p}_1 \\ \mathbf{p}_1 \end{pmatrix} = \begin{pmatrix} \alpha \mathbf{y}(v_l) \\ \beta \mathbf{y}(v_r) \end{pmatrix} \quad (10)$$

α and β are already determined by the two equations for \mathbf{p}_1 . For cubic Béziers (3) implies the conditions

$$\begin{pmatrix} \mathbf{p}_0 & \mathbf{y}(v_l) \\ \mathbf{p}_3 & \mathbf{y}(v_r) \end{pmatrix} \begin{pmatrix} \mathbf{p}_1 \\ \mathbf{p}_2 \end{pmatrix} = \begin{pmatrix} \alpha \mathbf{y}(v_l) \\ \beta \mathbf{y}(v_r) \end{pmatrix} \quad (11)$$

Thus we only have the two free (scalar) parameters α and β . With this notations we can formulate two strategies for curve approximation with cubic Béziers.

- 1) Determine α and β by the condition $\mathbf{y} \cdot \mathbf{v} = 0.5$ $\mathbf{x} = 0.5$ – interpolation in between.
- 2) Determine α and β by the condition $\sum_{i=1}^n \|\mathbf{y} \cdot \mathbf{v} - s_i\|^2 \rightarrow \min$ with $0 \leq s_1 \leq s_2 \leq \dots \leq s_n \leq 1$ – (discrete) mean square approximation. Here we only use $n = 3$ with $s_1 = 1/4, s_2 = 1/2$ and $s_3 = 3/4$.

We further use the error estimators

$$\delta_i = \max_{i=1,3} \left\| \mathbf{y} \cdot \mathbf{v} - \frac{i}{4} - \mathbf{x} \cdot \frac{i}{4} \right\|_2 \quad (12)$$

in 1) and

$$\delta_i = \max_{i=1,3,5,7} \left\| \mathbf{y} \cdot \mathbf{v} - \frac{i}{8} - \mathbf{x} \cdot \frac{i}{8} \right\|_2 \quad (13)$$

in 2). If the error is too large, we (recursively) subdivide v_l, v_r into two equal parts.

We use $\mathbf{y}_h: \mathbf{y} \cdot \mathbf{v} = 1/2$, $\mathbf{t}_0: \mathbf{y} \cdot \mathbf{v}_l$ and $\mathbf{t}_1: \mathbf{y} \cdot \mathbf{v}_r$. Strategy 1) results in an equation

$$A\mathbf{x} = \mathbf{b} \quad (14)$$

with

$$A = \begin{bmatrix} \frac{3}{8}\mathbf{t}_0 & \frac{3}{8}\mathbf{t}_1 \end{bmatrix} \quad (15)$$

$$\mathbf{b} = \mathbf{y}_h \cdot \begin{bmatrix} \frac{1}{2}\mathbf{p}_0 & \frac{1}{2}\mathbf{p}_3 \end{bmatrix} \quad (16)$$

and

$$\mathbf{x} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \quad (17)$$

With $\mathbf{z}_i: \mathbf{y} \cdot \mathbf{v} = s_i$ strategy 2) leads to an overdetermined linear system

$$A\mathbf{x} = \mathbf{b}, \quad \min \quad (18)$$

with

$$A = \begin{bmatrix} \frac{27}{64}\mathbf{t}_0 & \frac{9}{64}\mathbf{t}_1 \\ \frac{3}{8}\mathbf{t}_0 & \frac{3}{8}\mathbf{t}_1 \\ \frac{9}{64}\mathbf{t}_0 & \frac{27}{64}\mathbf{t}_1 \end{bmatrix} \quad (19)$$

$$\mathbf{b} = \begin{bmatrix} \mathbf{z}_1 \cdot \frac{27}{32}\mathbf{p}_0 & \frac{5}{32}\mathbf{p}_3 \\ \mathbf{z}_2 \cdot \frac{1}{2}\mathbf{p}_0 & \frac{1}{2}\mathbf{p}_3 \\ \mathbf{z}_3 \cdot \frac{5}{32}\mathbf{p}_0 & \frac{27}{32}\mathbf{p}_3 \end{bmatrix} \quad (20)$$

and

$$\mathbf{x} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \quad (21)$$

Figure 5 shows an example for the approximation

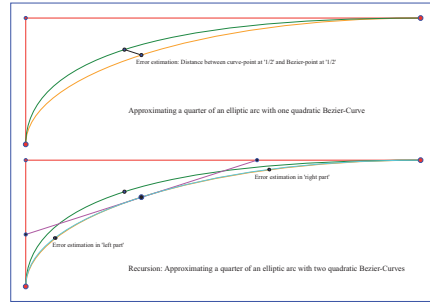


Figure 5: Recursive approximation with quadratic Bézier curves

tion of one quarter of an ellipse with quadratic Béziers. Only one level of recursion is shown. The error estimator – at least for higher recursion level – is rather pessimistic. This is due to the fact that the error vectors become more and more tangential to the curves with higher recursion level. For a measurement of the real curve to curve distance the error vectors should be orthogonal to them. We will fix this lack later on with parameter corrections. This simple example already shows the power of spline approximation. For quadratic splines we expect the convergence rate $\mathcal{O}(h^3)$. This can only be achieved if the underlying curve is in C^3 $v_{\min} \leq v_{\max}$. Thus for every recursion the error will be (asymptotically) reduced by a factor of 8. In figure 6 we have used a cubic Bézier with strategy 2) for a Lissajous figure with dimension 20×14 . We made two approximations. One with a desired absolute error of 10^{-1} and as reference one with an absolute error of 10^{-6} . Additionally the control polygon is shown. Very few control points lead to a good approximation of the Lissajous curve. For cubic splines we expect the convergence rate $\mathcal{O}(h^4)$. Here the curve for approximation has to be in

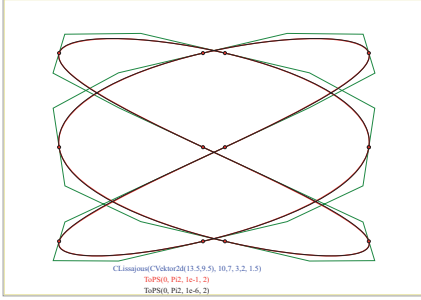


Figure 6: Approximation with cubic Bézier curves

$C^4 v_{min} v_{max}$.

Figure 7 shows the approximation and estimation strategy. Again we have use a quarter of an ellipse. Only the entrance level is shown. If we

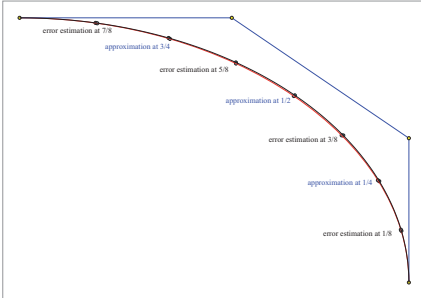


Figure 7: Approximation with cubic Bézier curves – strategy and error estimation

use the error estimation as shown in figure 5 and figure 7 with several recursion levels we will not recognize the expected convergence rates. This is due to the fact that the used error estimators do not measure the distance orthogonal the curves. To overcome this lack we have to correct the parameter values for v originally given as $v s_i$ according to (9) in such a way that we measure orthogonal. Using $\mathbf{p} \times \mathbf{s}_i$ this can be formulated

as

$$f(v) = \mathbf{y}(v)^T \mathbf{y}(v) - \mathbf{p}^T \mathbf{0} \quad (22)$$

The derivative of this function is

$$f'(v) = \mathbf{y}(v)^T \mathbf{y}(v) - \mathbf{p}^T \mathbf{y}(v)^T \mathbf{y}(v) \quad (23)$$

and it is easy to implement Newton's method for (22). For surfaces this works similarly. (22) is replaced by the system

$$f(u, v) = \begin{bmatrix} \mathbf{x}_u^T \mathbf{u} & \mathbf{x}_v^T \mathbf{u} & \mathbf{p} \\ \mathbf{x}_u^T \mathbf{v} & \mathbf{x}_v^T \mathbf{v} & \mathbf{p} \end{bmatrix} \quad (24)$$

and the derivative is given by (we omit the parameters and use $\mathbf{d} = \mathbf{x}_u \mathbf{v} - \mathbf{p}$)

$$f' = \begin{bmatrix} \mathbf{x}_{uu}^T \mathbf{d} & \mathbf{x}_{uv}^T \mathbf{x}_u & \mathbf{x}_{uv}^T \mathbf{d} & \mathbf{x}_{uv}^T \mathbf{x}_v \\ \mathbf{x}_{uv}^T \mathbf{d} & \mathbf{x}_{uu}^T \mathbf{x}_v & \mathbf{x}_{vv}^T \mathbf{d} & \mathbf{x}_{vv}^T \mathbf{x}_v \end{bmatrix} \quad (25)$$

This parameter correction already leads to an error estimator that reflects the expected convergence rates.

The results can be improved if we iterate according to (18) – (21) with the corrected v 's and thus modified \mathbf{z}_1 , \mathbf{z}_2 and \mathbf{z}_3 in (20). This decoupled iteration only slowly converges towards the optimal solution. This is due to the fact that approximation (18) minimizes a point distance and not the curve distance.

In [8] and [9] an approach based on the minimization of a quadratic approximant of the squared distance function was introduced. Unfortunately in general the second order Taylor approximant does not lead to symmetric positive definite system matrices and for this reason we can not guarantee the existence and uniqueness of the minimum. For this reason the second order Taylor approximation was modified to guarantee positive definiteness. But this modification destroys the second order and thus the claimed quadratic convergence of that method. Nevertheless such methods need less iterations. But the drawback is a computational overhang. For instance the curvature – the principal curvature and directions for surfaces – has / have to be computed. Furthermore the parametrization is not

really totally avoided. In [4] the behaviour of the point distance minimization (PDM) and the squared distance function minimization (SDM) was studied. Regarding run time the authors conclude: *Our experiments show that the time used by SDM on iterative optimization is about 30% to 50% more than PDM because SDM needs extra time to set up the more complex SDM error function.*

We have used a different approach. This avoids the curvature computation. Furthermore our method can still be written in the form of (18) which is not the case for SDM. We will demonstrate the idea of our method for the approximation with B-splines. Due to (5) and (6) B-spline curve or surface points are linear combinations of the control points.

$$\mathbf{x}_i = \mathbf{x}_{t_i} \sum_j a_{ij} \mathbf{p}_j \quad (26)$$

The \mathbf{x}_i correspond to curve points $\mathbf{y}_i = \mathbf{y}_{v_i}$ and we have to minimize the distances $\|\mathbf{x}_i - \mathbf{y}_i\|_2$. We collect the control points \mathbf{p}_j in a vector \mathbf{p} , the \mathbf{y}_i in a vector \mathbf{y} and coefficients a_{ij} in a sparse matrix A . Now PDM is simply

$$\|A\mathbf{p} - \mathbf{y}\|_2 \min \quad (27)$$

For the solution of (27) we can use orthogonal transformations or the normal equations. For more details especially on surfaces and volumes see ([1]). To reduce the number of parameter corrections (the *outer loop*) we minimize the distances of the \mathbf{y}_i to the linear approximation of \mathbf{y}_{v_i} at the point \mathbf{y}_i . Note that this coincides with the squared distance function for points on the curve. Thus we have to minimize

$$\mathbf{n}_i^T \mathbf{x}_i - \mathbf{y}_i \quad (28)$$

where $\mathbf{n}_i = \mathbf{n}_{v_i}$ is orthogonal to \mathbf{y}_{v_i} and normalized. Notice that (28) is a minimization over the control points \mathbf{p}_j . To use matrix notations we have to separate the x and y components. We define $N_x = \text{diag } \mathbf{n}_{i,x}$, $N_y = \text{diag } \mathbf{n}_{i,y}$ and collect

the terms $\mathbf{n}_i^T \mathbf{y}_i$ in a vector \mathbf{d} . Now we can write this part as

$$\|N_x A \mathbf{p}_x - N_y A \mathbf{p}_y - \mathbf{d}\|_2 \min \quad (29)$$

For our final minimization problem we use (29) and a scaled portion of (27). Here we have to split \mathbf{p} into its x -components \mathbf{p}_x and y -components \mathbf{p}_y . The same way we split \mathbf{y} into \mathbf{y}_x and \mathbf{y}_y . With

$$\hat{A} = \begin{pmatrix} \lambda A & 0 \\ 0 & \lambda A \\ N_x A & N_y A \end{pmatrix} \quad (30)$$

$$\hat{\mathbf{p}} = \begin{pmatrix} \mathbf{p}_x \\ \mathbf{p}_y \end{pmatrix} \quad (31)$$

and

$$\hat{\mathbf{y}} = \begin{pmatrix} \lambda \mathbf{y}_x \\ \lambda \mathbf{y}_y \\ \mathbf{d} \end{pmatrix} \quad (32)$$

our minimization problem now reads.

$$\|\hat{A} \hat{\mathbf{p}} - \hat{\mathbf{y}}\|_2 \min \quad (33)$$

Thus again we can use orthogonal transformations, the normal equations or iterative methods. Especially for recursive approximation of surfaces and volumes the iterative methods are much faster. If only equidistant knot vectors are used a matrix free implementation is possible. The parameter $\lambda = 0$ is chosen depending on the error estimator. For small errors we use smaller λ s. If the oversampling is high enough we even choose $\lambda = 0$ for high accurate approximations.

As iterative solver we normally use CGLS, a Conjugate Gradient method for linear Least Squares (also called CGNR in [10]). To apply CGLS we only need an effective implementation to multiply the system matrix A and its transpose A^T with vectors. It should be noted that in our applications A is very sparse and for recursive strategies we have very good starting vectors. For further informations and details on the use of CGLS see [10] or [3].

4. OBJECT RECOGNITION

Sometimes it is useful to have tools that recognize given curves as lines, circles or conics. We have implemented that in *WinCAG* in several ways. In this section we will report on the module that generates data polygons by free-hand mouse-drawing and then determines the best fit from a given list of objects. An example is shown in figure 1. The three recognized objects are a line, a circle and an ellipse. Hyperbolas and parabolas will be recognized too. For this purpose the user has to give tolerances for each object. Since for instance an ellipse normally fits better than a circle, the tolerance for circles must be chosen larger than that for ellipses. The system stores the data polygons and the user can force it to compute a given kind of curve. The curves corresponding to the data polygons $(x_i, y_i) \quad i = 0 \dots N$ are computed by least squares solutions. For this reason we start with the implicit form of lines, circles and conics:

$$\begin{aligned} ax + by + c &= 0 \\ x^2 + y^2 - r^2 &= 0 \\ \sum_{i,j=0}^2 a_{ij} x^i y^j &= 0 \end{aligned} \quad (34)$$

For circles the equations seem to be non-linear. But

$$\begin{aligned} x^2 + y^2 - r^2 &= 0 \\ 2xx_m - 2yy_m + x_m^2 - y_m^2 - r^2 &= 0 \\ 2xx_m - 2yy_m + \alpha &= 0 \end{aligned} \quad (35)$$

with the substitution $\alpha = x_m^2 - y_m^2 - r^2$ it is linear in x_m, y_m and α and has the same solution as the original system. If we plug in the (x_i, y_i) we get a linear least squares problem for the coefficients in all cases. For lines and conics the coefficients have an arbitrary non zero scaling factor. We can choose one of them to be one. In very few cases the chosen coefficient must be zero and we have to revise our guess and have to solve the system once more. The principal axis transformation is used to get the kind of the

conic. The conic fit by the implicit form does not yield the optimal (non-linear) least squares solution given by the parameterized form. Thus we determine the parameters for the data points and make some Gauss-Newton steps to improve our result. If we are *close to* a parabola the system computes a least squares solution for this case, too. Then the user can optionally decide to take the parabola.

If the distances to all of the above objects is too large the system will compute a B-spline approximation. Figure 3 shows an example for that situation. Details on spline approximation have been given in the previous section. The only difference here the determination of the parametrization of the knot vector. We use a chord length parametrization since it is known that in general this leads to small approximation errors. We refer to [6]

If we want to force the system to compute a special conic section we have to make some extra efforts for parabolas and in the case the principal axis transformation does not lead to the right kind of the conic. Parabolas do not cause a real problem. In general the implicit solution leads to feasible initial values for the non linear iteration.

Next we describe the approach for forcing an elliptic fit. With the *normal* form

$$\mathbf{x}_n(t) = \begin{pmatrix} a \cos t \\ b \sin t \end{pmatrix} \quad (36)$$

and the rotation matrix

$$D(\psi) = \begin{pmatrix} \cos \psi & \sin \psi \\ \sin \psi & \cos \psi \end{pmatrix} \quad (37)$$

the general form of an ellipse is given by

$$\mathbf{x}(t) = \begin{pmatrix} x_m \\ y_m \end{pmatrix} + D(\psi) \mathbf{x}_n(t) \quad (38)$$

We have five parameters a, b, x_m, y_m and ψ and denote the corresponding Ellipse by $E(a, b, x_m, y_m, \psi)$. Given a point \mathbf{x}_i and the parameters a, b, x_m, y_m and ψ we can compute

the distance d_i \mathbf{x}_i ; a b x_m y_m ψ from the point \mathbf{x}_i to the ellipse E a b x_m y_m ψ . For a point cloud \mathbf{x}_i $i = 1 \dots N$ we can define the distance function as (in d_i we have omitted the arguments)

$$\text{dist } a \ b \ x_m \ y_m \ \psi = \sum_{i=1}^N d_i \quad (39)$$

This function is highly non linear. For the minimization of $\text{dist } a \ b \ x_m \ y_m \ \psi$ we use the Nelder-Mead-method published in [7]. This method has only linear convergence but is very robust and does not need any derivatives. As starting values we use $\psi = 0$, $a = b = r$ with r and x_m and y_m from a circle approximations. In our examples this method never fails.

5. RECONSTRUCTION FROM IMAGES

For the reconstruction of curves from images we only assume that the input format can be transferred into RGB-format. In a first step we cluster the contained colors in the 3D RGB color space and compute local color centroids. The pixels of each centroid are handled separately looping over all centroids. Starting with a low number of neighbored points we use the methods from the previous section to compute initial objects of different kinds. In a next step we insert further points that have a short distance to the previously computed objects and recompute the minimization problems with more points. If no further points close the actual objects are found and the extension is not negligible we store the best fit and drop the used points from our list. If at the end of this process a significant number of points is remaining we try to use B-spline fits with standard methods for point clouds to get connect curve parts. This may fail for some points and we have to leave this points untreated. In our examples this normally only happens for noisy pixels that do not belong to any curve part. The methods only fail for scenarios with a very large number of intersections or very close curves. This is due to the fact that we remove the used points and for the last curves in this cases we do not have enough points remaining for the curve

reconstruction.

6. CONCLUSIONS

We have demonstrated the potential of effective approximations with Bézier or B-spline curves applied on scattered data or the conversion of curves given in other formats. One of the main aspects is the preparation of websites and / or further use in a CAGD system. The major advantage of the used vector graphics formats is the very small size of the files and the possibility to zoom without grid pattern effects. This leads to the usability on tablets and smartphones and the potential to make significant improvements in how geometry is learnt and taught.

INTERNET SOURCES ON VECTOR GRAPHICS

- PostScript Adobe: http://partners.adobe.com/public/developer/ps/index_specs.html
- Wikipedia-PostScript: <http://de.wikipedia.org/wiki/PostScript>
- Scalable Vector Graphics: <http://www.w3.org/TR/SVG11/>
- The authors website on this topic: <http://www.igpm.rwth-aachen.de/brakhage/ICGG2014>

REFERENCES

- [1] K.-H. Brakhage. High quality mesh generation and sparse representation using b-splines. In B.K. Soni. et al, editor, *7th International Conference on Numerical Grid Generation in Computational Field Simulations*. Resort Whistler, British Columbia, Canada, September 25-28 2000.
- [2] K.-H. Brakhage. WinCAG - dynamical geometry for teaching and learning. In *13th International Conference on Engineering Computer Graphics and Descriptive Geometry*. Dresden, Germany, August 3-8 2008.

- [3] K.-H. Brakhage. Grid generation and grid conversion by subdivision schemes. In B.K. Soni et al, editor, *11th International Conference on Numerical Grid Generation in Computational Field Simulations*. Montreal, Canada, May 24-28 2009.
- [4] K. Cheng, W. Wang, H. Qin, K.-Y. Wong, H. Yang, and Y. Liu. Fitting subdivision surfaces to unorganized point data using sdm. In *In Pacific Conference on Computer Graphics and Applications 2004*, pages 16–24. 2004.
- [5] G. Farin. *Curves and Surfaces for CAGD. A Practical Guide*. The Morgan Kaufmann Series in Computer Graphics and Geometric Modeling, fifth edition, 2002.
- [6] M. Floater. Arc length estimation and the convergence of polynomial curve interpolation. *BIT Numer. Math.*, 45: 679–694, 2005.
- [7] J. Nelder and R. Mead. A simplex method for function minimization. *The Computer Journal*, 7(4): 308–313, 1965.
- [8] H. Pottmann and M. Hofer. Geometry of the squared distance function to curves and surfaces. In *VISUALIZATION AND MATHEMATICS III*, pages 223–244. 2003.
- [9] H. Pottmann and S. Leopoldseder. A concept for parametric surface fitting which avoids the parametrization problem. *Computer Aided Geometric Design*, 20: 343–362, 2003.
- [10] Y. Saad. *Iterative Methods for Sparse Linear Systems*. SIAM, second edition, 2003.

research interests are Computer Aided Geometric Design, CAx Technologies, Grid Generation, Scientific Computing, Computer Graphics, and Development of Education Software. He can be reached by e-mail: brakhage@igpm.rwth-aachen.de, by Fax: +49(241)8092317, by phone: +49(241)8096591, the postal address: Inst. of Geometry and Numerical Mathematics / RWTH Aachen / Templergraben 55 / D-52056 Aachen, Germany, or through the website: <http://www.igpm.rwth-aachen.de/brakhage>.

ABOUT THE AUTHOR

Karl-Heinz Brakhage is member of the Institute of Geometry and Numerical Mathematics at the Aachen University of Technology. His

THE EDUCATIONAL RESEARCH PROJECT GEODIKON POINTING ACCURACY, STRATEGIES AND GENDER-SPECIFIC EFFECTS

Günter MARESCH

University of Education Salzburg, Austria

ABSTRACT: The article discusses the results of the analyses of the pretests of the spatial ability project GeodiKon. 903 students were taking part in this project in which the major aim was to find out whether the training of each factor of spatial ability and of a repertoire of strategies for solving spatial tasks will lead to an improvement of spatial ability. It will refer to the findings regarding (1) the pointing accuracy as a function of pointing direction, (2) the usage of different strategies for solving spatial tasks, (3) promising strategies for solving spatial tasks, and (4) gender-specific effects.

Keywords: spatial ability, pointing accuracy, strategies for solving spatial tasks, gender-specific effects, visualization, spatial perception, spatial relations, mental rotation, spatial orientation.

1. THE RESEARCH QUESTION

The research project GeodiKon (development of a didactical concept for geometry education) has been accepted for funding by the Austrian Federal Ministry for Education, the Arts and Culture and the University of Education Salzburg (October 2012). The project deals with the question of optimal support for and the development of the factors of spatial ability, and also with the deliberate training of different strategies for solving spatial tasks. Altogether 46 test classes with 903 students at the age of 12 to 14 years from all types of secondary schools (Hauptschule (HS), Neue Mittelschule (NMS), Bundesrealgymnasium (BRG), and Bundesgymnasium (BG)) in the three Austrian provinces of Salzburg, Styria, and Lower Austria have been taking part in the project in the years 2013 and 2014.

The underlying hypothesis of the project is the following: Training (making aware, categorising, internalizing) of each factor of spatial ability and training of the repertoire of strategies for solving spatial tasks will lead to an im-

provement of spatial ability.

The major aims of the project are:

- Development of specific (paper and pencil) learning materials for the training of the four factors visualization, spatial relations, mental rotation, and spatial orientation of spatial ability [11, 12, 13, 19] with the aim to train a balanced and extensive development of the learners' spatial ability
- Development of a structured model of strategies for solving spatial tasks [1, 8, 14, 15, 17] with the aim to extend the learners' repertoire of problem solving strategies
- Providing a user-friendly guideline with all specific learning materials of the project for lecturers for mathematics and geometry
- Training for teachers and lecturers how to use the materials in classes
- Dissemination of the results of the project in conference presentations and papers.

2. THE PROGRESS OF THE PROJECT

During the first phase of the project (January until September 2013) the Austrian project team (headed by G. Maresch, including mem-

bers from the University of Education of Salzburg, the University of Education of Lower Austria, the University of Education of Styria, the University of Education of Vienna, the University of Innsbruck, the University of Salzburg, the University of Vienna, the Technical University of Vienna, and the workgroup for didactical innovation) compiled learning materials for 12 weeks of lessons in geometry and mathematics. The learning material contains specific spatial ability tasks training the factors of spatial ability and the different strategies for solving spatial tasks. The structured model of the four pairs of strategies (holistic strategy, analytical strategy, spatial thinking, planar thinking, move object, move self, verifying strategy and falsifying strategy) for the solution of spatial tasks was developed [14, 15] and the test battery was set up with the support of the University of Vienna and the Technical University of Vienna.

The second phase of the project (September 2013 until February 2014) focussed on the pretests and posttests and the learning phases in classes.

The 46 test classes were divided into four project groups:

- Group A: The students of the group-A-classes get information about strategies for the solution of spatial tasks and they work with the newly developed learning materials
- Group B: The students of the group-B-classes work with the newly developed learning materials (but get no information about strategies)
- Group C: The students of the group-C-classes are students who are taught geometry as a separate subject in their school curriculum for at least between one to two lessons per week (and they do not work with the developed learning materials and get no information about strategies for the solution of spatial tasks)
- Group D: The students of the group-D-classes are students of the control group (and they do not work with the developed learning materials and get no information about strategies and they even have no separate subject geometry in

their school curriculum).

In September and October 2013 the pretests were carried out. Directly after the pretests the twelve-week long learning phase began for all groups. In January and February 2014 the posttests were carried out in all the 46 school classes. The third phase (March until December 2014) of the project started after the posttests: digitalization, preparation and analysis of the collected data, compilation of the user-friendly book with all the special learning materials of the project, and dissemination of the results of the project in conference presentations and papers.

3. THE TEST BATTERY

The test battery of the pretest and the posttest consisted of four spatial ability tests (Three Dimensional Cube Test (3DW) [4], Differential Aptitude Test (DAT) [2], Mental Rotation Test (MRT) [16] and Spatial Orientation Test (SOT) [6] and additional questions about the spatial tasks used for solving strategies and questions about age, gender and interests. The pretest lasted for 85 minutes and the posttest for 77 minutes. For better understanding of the pretest results of the project, the four spatial ability tests which we used are explained as follows:

3.1 Three Dimensional Cube Test (3DW) [4]

This test investigates whether the cube A, B, C, D, E or F is exactly the same cube as the cube X or whether the right answer is G (no cube matches). If a student does not know the solution, she/he has to choose H ("I do not know the answer"). Each pattern at the side faces of the cube occurs only once.

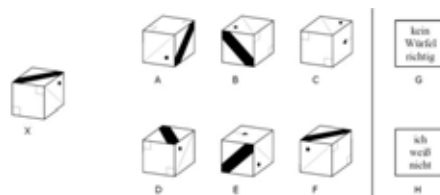


Figure 1: An example of the Three Dimensional Cube Test (3DW) [4]

3.2 Differential Aptitude Test (DAT) [2]

The tasks of this test consist of handling folding templates with shades and patterns. The templates can be folded to three dimensional objects. Each task shows one folding template and four three dimensional objects. Students have to choose which of these three dimensional objects A, B, C or D can be folded from the folding template provided.

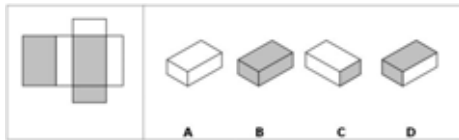


Figure 2: An example of the Differential Aptitude Test (DAT) [2]

3.3 Mental Rotation Test (MRT) [16]

An object is presented on the left. The probands have to determine which two of the four sample stimuli on the right are rotated versions of the target stimulus [16].

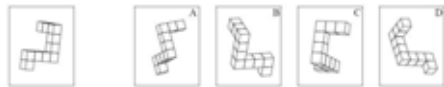


Figure 3: An example from the Mental Rotation Test [16]

3.4 Spatial Orientation Test (SOT) [6]

This is a test of one's ability to imagine different perspectives or orientations in space. In each task one can see a picture of an array of objects and an "arrow circle" with a question about the direction between some of the objects. For each question one should imagine to be standing next to one object in the array (which is named in the center of the circle) and facing another object, named at the top of the circle. The task is to draw an arrow from the center object showing the direction to a third object

from this facing orientation [10].

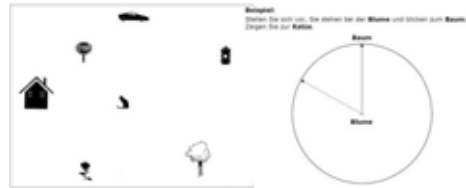


Figure 4: An example of the Spatial Orientation Test [6]

4. RESULTS OF THE PRETESTS

46 test classes with 903 students at the age of 12 to 14 years from all types of secondary schools (Hauptschule (HS), Neue Mittelschule (NMS), Bundesrealgymnasium (BRG), and Bundesgymnasium (BG)) in the three Austrian provinces of Salzburg, Styria and Lower Austria took part in the pretests in September and October 2013. The statistical analysis was carried out by Erich Svecnik [18]. In the following we will present four of the results.

4.1 Sensitivity of Gender, Type of School and Level of Education

16.3% of the variance in the 3DW Test and 11.8% in the DAT can be explained against the background variables gender, school type and school level. There are no gender differences in the 3DW Test and the DAT. The results of the MRT can be explained with 19.5% regarding gender, school type and school level. At the MRT male students attained about 1.83 points more than female students. Male students managed to solve 9.82 items correctly whereas female students had an average solving rate of 7.27 items. We can nearly see the same gender difference when we study the processing speed – although in a moderated form: Male students edited 11.31 items (not taking account of the correctness of the solution) and female students edited 10.90 items ($F=9.24$; $p=0.002$).

In the SOT the average deviation is 59.04%. Male students had a lower average deviation of

14.4° than female students (male 52.6% and female 67%).

The analysis of the pretest results of the project is in accordance with current scientific literature [5], which points out that the gender differences regarding spatial ability are not as extensive as assumed in former times, and that differences can be observed only in the two factors mental rotation (mainly in tests under time pressure) and spatial orientation.

4.2 Which Solution Strategies for Spatial Tasks Are Used by Students?

After each of the four spatial tests the students once again got one of the tasks, which was arbitrarily chosen. When the students solved the task they were asked to observe themselves accurately with which spatial strategy they solved the task. Then students answered four questions concerning the different strategies which they used (four pairs of strategies) – each in an eight-part scale:

1a: Looked at the object in its entirety (whole approach – holistic strategy): You looked at the whole object. You did not concentrate on parts of the object only. You imagined the whole object and found the solution right away.

1b: Looked at parts of the object (part approach – analytic strategy): You concentrated on parts of the object only. You did not have to use the whole object for the solving process.

2a: Spatial thinking: You created a mental, three-dimensional model of the object and solved the task by working on this mental model.

2b: Planar thinking: You created a mental, planar (two-dimensional) model of the setting and solved the task by working on this planar mental model.

3a: Move self: You placed yourself inside the setting and moved around mentally and changed your perspective.

3b: Move object: You positioned yourself mentally as an observer outside the setting and moved (rotated, moved,...) the individual objects.

4a: Falsifying strategy: You identified all the incorrect solutions first and excluded them step by step.

4b: Verifying strategy: You had the correct solution in mind and worked on it directly.

These questions were answered in a differentiated and satisfying way. So this can be seen as an indication that for the solving process of the items of all four tests students really used the strategies which were postulated theoretically. In particular, we can see that students used the move self strategy very often in the SOT as compared to the other three tests in which mainly the move object strategy was used. As expected, we can also ascertain that students used the planar thinking strategy more often in the SOT than in the other three tests. We also see that students used the holistic strategy much more than the analytical strategy in the MRT. In general, the probands used similar strategies in the DAT and the 3DW Test.

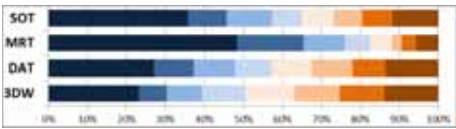


Figure 5a: Whole approach – Holistic strategy (left side) vs. Part approach – Analytical strategy (right side)

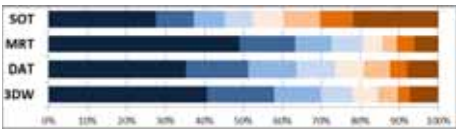


Figure 5b: Spatial thinking (left side) vs. Planar thinking (right side)

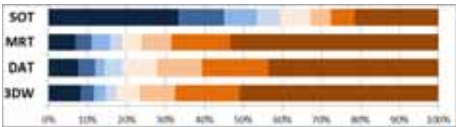


Figure 5c: Move self (left side) vs. Move object (right side)

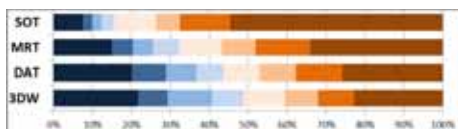


Figure 5d: Falsifying strategy (left side) vs. Verifying strategy (right side)

When, as a first step, we analyzed the answers concerning the strategies we were interested in the question: Can we identify gender differences?

3DW Test: Gender differences can only be identified in the fourth pair of solution strategies (falsifying strategy - verifying strategy). Female students used the falsifying strategy much more often than male students ($F=6.92$; $p=0.009$).

DAT: The only gender difference which can be identified in the DAT is the same as in the 3DW Test. Female students used the falsifying strategy much more often than male students ($F=17.64$; $p<0.001$).

MRT: Even in the MRT female students used the falsifying strategy more often than male students ($F=8.72$; $p=0.003$). Female students did not use the holistic strategy as often as male students ($F=15.85$; $p<0.001$).

SOT: To use the falsifying strategy in this test seems to be hard to imagine. Therefore it is used very rarely. Nevertheless, in the SOT it was used more often by female students ($F=4.43$; $p=0.036$).

4.3 Do promising strategies for solving spatial tasks exist?

To find out more about the influence of the spatial task solution strategies used by the students on their performance, we used regression models in which gender, school type, school level and all the items of the strategy questions were included. Their contribution to the variance explanation has been determined beforehand.

In the 3DW Test the variance explanation increases from 16.3% to 18.0%, so the strate-

gies contributed significantly ($F=4.73$; $p=0.001$), most of all the two strategies analytic strategy ($\beta=0.096$; $p=0.005$) and the spatial thinking strategy ($\beta=0.078$; $p=0.025$).

In the DAT the variance explanation increases, if we include the spatial task solution strategies, from 11.8% to 15.6% ($F=9.48$; $p<0.001$). Just as in the 3DW Test the same two strategies contributed to the increase of the performance: The analytic strategy ($\beta=0.109$; $p=0.002$) and the spatial thinking strategy ($\beta=0.169$; $p<0.001$).

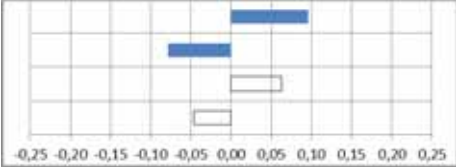
In the MRT the variance explanation increases from 16.4% to 23.7% ($F=18.68$; $p<0.001$). Three strategies contributed significantly: The holistic strategy ($\beta=0.115$; $p=0.001$) in contrast to the analytic strategy, the spatial strategy ($\beta=0.202$; $p<0.001$) in contrast to the planar strategy and the move object strategy ($\beta=0.098$; $p=0.003$) in contrast to the move self strategy. Only the falsifying strategy does not contribute significantly ($p=0.438$).

Finally, in the SOT the variance explanation increases from 15.7% to 18.7% ($F=7.58$; $p<0.001$), caused by the move object strategy ($\beta=0.093$; $p=0.007$) and the falsifying strategy ($\beta=0.150$; $p<0.001$).

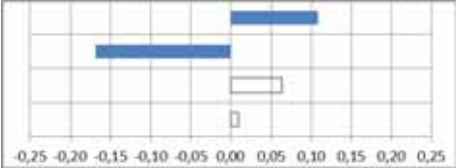
The question whether there exist promising strategies for solving spatial tasks must be answered in a differentiating manner. We can see that the same two promising strategies (analytical strategy and spatial strategy) are used in the 3DW Test and in the DAT. In contrast, in the MRT and the SOT we can notice that other strategies seem to be more promising (MRT: holistic strategy, spatial strategy, and move object strategy; SOT: move object strategy and the falsifying strategy). In figure 6 the most successful strategies for each of the four spatial tests 3DW Tests, DAT, MRT, and SOT are visualized. Statistically significant beta values are shown as filled bars, non-significant values as frames only. In the respective models the effects of gender, school type, and school level were considered, so that we can see the pure statistical effects in the figures.

First row: Holistic approach (left) – Analytic approach (right)
 Second row: Spatial thinking (left) – Planar thinking (right)
 Third row: Move self (left) – Move object (right)
 Forth row: Falsifying strategy (left) – Verifying strategy (right)

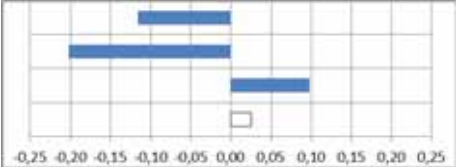
3DW Test:



DAT:



MRT:



SOT:

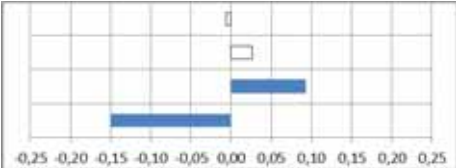


Figure 6: Promising strategies for spatial tasks. Statistically significant beta values are shown as filled bars, non-significant values as frames only [18]

4.4 Results of the Spatial Orientation Test

(SOT)

Average deviation:

In the SOT we can see that the performance of the 12 years old and 14 years old students is much lower (average deviation of 59.04°) than the performance of 17 years old students (average deviation of 30°) [3]. We assume that the factor spatial orientation of individuals is developed later than other factors of spatial intelligence.

Pointing accuracy as a function of imagined heading:

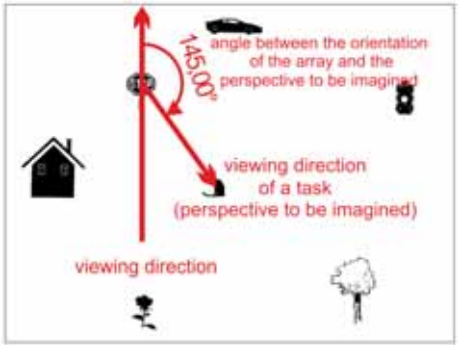


Figure 7: Angle between the orientation of the array and the perspective to be imagined at tasks of the SOT

We analyzed the assumption whether the absolute angular error increases with the angular deviation of one’s imagined heading (perspective) from the orientation of the array [10]. Figure 7 shows the angle between the orientation of the array and the perspective to be imagined at tasks of the SOT. At figure 8 we can see clearly that the absolute angular error does increase with the angular deviation of one’s imagined heading (perspective) from the orientation of the array in general. We therefore can confirm the results of Kozhevnikov & Hegarty [10].

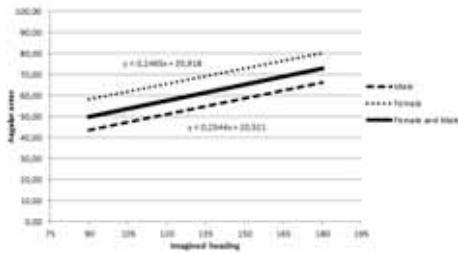


Figure 8: Pointing accuracy as a function of imagined heading

Pointing accuracy as a function of pointing direction:

Consistent with the research results of Kozhevnikov & Hegarty [10] we can confirm that pointing accuracies for front (0 degree deviation) and back (180 degree deviation) responses were significantly higher than those for right (90 degree deviation) and left (270 degree deviation) responses (see figure 9). Again, in all tasks male students were better than female students. Where the mean deviation for all students was 59.04°, female students had a deviation of 67° and male students of 52.6°.

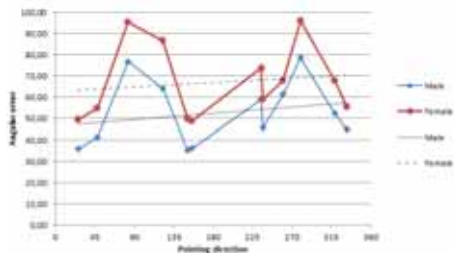


Figure 9: Pointing accuracy as a function of pointing direction

5. CONCLUSIONS

Test Battery and factors of spatial ability: The analysis of the results of the pretests of the project GeodiKon gave evidence that the test battery can be argued as being scientifically cor-

rect. The proportion of the common variance of the four tests (3DW Test, DAT, MRT, and SOT) is between 13% and 26%. So the tests do rarely measure common characteristics but their purpose is to measure special characteristics. It would be interesting to find out whether the tests used are really markers [7] of underlying factors of spatial ability, and whether it can be ascertained how much the four factors visualization, spatial relation, mental rotation, and spatial orientation - which have been used - have in common and how selective they are.

Gender specific effects: The analysis of the results of the pretests confirms current scientific literature [5], which points to the fact that gender differences as concerns spatial ability are not as distinct as was supposed in the past, and that the differences are only manifest in two of four factors of spatial ability (visualization, spatial relations, mental rotation, and spatial orientation): mental rotation (and here mainly in mental rotation tests under time pressure - e.g. MRT) and spatial orientation (SOT). We cannot see any gender differences in the 3DW Test and the DAT.

Spatial Orientation Test and pointing accuracy:

In the SOT we can see that the performance of the 12 to 14 years old students is much lower (average deviation of 59.04°) than the performance of 17 years old students (average deviation of 30°) [3]. It would be desirable to verify the assumption that the factor spatial orientation of individuals is developed later than other factors of spatial intelligence. A long term study would bring us more results. In accordance with Kozhevnikov & Hegarty [10] we can confirm that pointing accuracies for front and back responses were significantly higher than those for nearly right or left responses. Can we recognize this effect only in the SOT, or can we also find similar results in other paper pencil tests for spatial orientation, or even in tests for spatial ability in dynamic scenes on a screen, or even in nature?

Strategies for solving spatial tasks: All students answered the questions pertaining to the spatial tasks solving strategies used in a differentiated and sufficient manner. So this is an indication that for the solving process of the items of all four tests students really used the strategies which were postulated only in theory. We can also claim that students really used different strategies for different tests. It can also be stated that female and male students used different strategies in the same tasks, e.g. in the 3DW Test and the DAT. Female students much more often used the falsifying strategy than male students.

The underlying hypothesis of the project GeodiKon is: Training (making aware, categorising, internalizing) of each factor of spatial ability and training of the repertoire of strategies for solving spatial tasks will lead to an improvement of spatial ability. After the posttests we will therefore analyze all the data to find out whether the knowledge of strategies for solving spatial tasks will improve spatial abilities or not, or if there will have to be a more complex answer to the question.

ACKNOWLEDGMENTS

This is part of the research project GeodiKon, funded by the Federal Ministry for Education, the Arts and Culture (BMUKK-20.040/0012-I/7/2012) and the University of Education of Salzburg. Project partners come from the University of Education of Lower Austria, the University of Education of Styria, the University of Education of Vienna, the University of Innsbruck, the University of Salzburg, the University of Vienna, and the Technical University of Vienna.

REFERENCES

- [1] B. S. Barratt. An analysis of verbal reports of solving problems as an aid in defining spatial factors. In: *The Journal of Psychology*, 36, 1953.
- [2] G. K. Bennett, H. G. Seashore, and A. G. Wesman. *Differential aptitude tests, forms S and T*. New York: The Psychological Corporation, 1973.
- [3] A. Duenser. *Trainierbarkeit der Raumvorstellung mit Augmented Reality*. Universität Wien, Fakultät für Psychologie, Dissertation, 2005.
- [4] G. Gittler. Entwicklung und Erprobung eines neuen Testinstruments zur Messung des räumlichen Vorstellungsvermögens. In: *Zeitschrift für Differentielle und Diagnostische Psychologie*, 2, 141-165, 1984.
- [5] J. Glueck, H. Kaufmann, A. Duenser, K. Steinbuegl. Geometrie und Raumvorstellung – Psychologische Perspektiven. In: *Informationsblätter der Geometrie* (IBDG) 24/1, 4-11, 2005.
- [6] M. Hegarty and D. Waller. A dissociation between mental rotation and perspective-taking spatial abilities. In: *Intelligence*, 32, 175-191, 2004.
- [7] M. Hegarty and D. Waller. Individual differences in spatial abilities. In P. Shah and A. Miyake (eds.): *The Cambridge Handbook of Visuospatial Thinking*, Cambridge University Press, 121-169, 2005.
- [8] M. A. Just and P. A. Carpenter. Cognitive Coordinate Systems: Accounts of Mental Rotation and Individual Differences in Spatial Ability. In: *Psychological Review*, 92, 1985.
- [9] P. Kline. *The handbook of psychological testing* (2nd ed.). London: Routledge, 1999.
- [10] M. Kozhevnikov, M. Hegarty. A dissociation between object-manipulation spatial ability and spatial orientation abilities. In: *Memory and Cognition* 29, 745-756, 2001.
- [11] M. C. Linn, A. C. Petersen. Emergence and characterization of sex differences on spatial ability: a meta-analysis. In: *Child Development*, 56, 1479-1498, 1985.

- [12]H. P. Maier. Räumliches Vorstellungsvermögen: Komponenten, geschlechtsspezifische Differenzen, Relevanz, Entwicklung und Realisierung in der Realschule. In: *Europäische Hochschulschriften*, Reihe 6, Psychologie, Band 493, 1994.
- [13]G. Maresch. Spatial Ability – The Phases of Spatial Ability Research. In *Journal for Geometry and Graphics*, Helder mann, 2014.
- [14]G. Maresch. Strategies for Assessing Spatial Ability Tasks. In: *Journal for Geometry and Graphics*, Helder mann, 2014.
- [15]G. Maresch. Strategien für die Bearbeitung von Raumvorstellungsaufgaben. In: *Informationsblätter der Geometrie (IBDG)*, Jahrgang 33, Heft 1, 2014
- [16]M. Peters, B. Laeng, K. Latham, M. Jackson, R. Zaiyouna, and C. Richardson. A Redrawn Vandenberg & Kuse Mental Rotations Test: Different Versions and Factors that affect Performance. In: *Brain and Cognition*, 28, 39-58, 1995.
- [17]K. Schultz. The contribution of solution strategy to spatial performance. In: *Canadian Journal of Psychology*, 45, 1991.
- [18]E. Svecnik. *GeodiKon Pretest - Ergebnisse der Datenanalysen*, PH Salzburg, 2013.
- [19]L. L. Thurstone. Some primary abilities in visual thinking. In: *Psychometric Laboratory Research Report*. 59, University of Chicago Press, Chicago, 1950.

ABOUT THE AUTHOR

Günter Maresch is a researcher at the University of Education Salzburg (Austria) and at the University of Salzburg (Austria). He has studied mathematics and descriptive geometry at the Technical University of Vienna (Austria) and is lecturer at the University of Salzburg and spread his concepts of elearning and computer aided design even to South Africa. See also www.geotic.at.

AN EDUCATIONAL TRIAL ON DRAWING FOR TRUE SHAPE BY USING AUXILIARY VIEW METHOD AIDED WITH CAD TO THE STUDENTS WITH HEARING IMPAIRMENTS

Iwao HONMA¹ and Kazunori KUWABARA²

¹Tsukuba University of Technology, Japan ²Hokkaido High School for the Deaf, Japan

ABSTRACT: It seems that drawing with auxiliary view method is difficult for students according to our former research [1]. Especially for the students with hearing impairments whose spatial ability is not high, it is more difficult to teach its method. This time, we tried to them to understand it by rather practical way. It was a way to make simple solid model with paper by making additional drawing for the surface which is not shown in the given drawing. By this trial, we got some important effects as follows. Aiming at completion of model, some students wrestled with the understanding of spatial relation, and at last they seemed able to understand it, and then, they reached the correct drawing and modeling. This result thought to be an encouraging educational effect which we did not expect. It seemed rather effective for those students to put them in such situation as that they must seek and try to draw true shapes by themselves, which were not shown on the drawing, or they could not complete modeling. In this process, it seems that the needs to complete the model helped their realistically thinking.

Keywords: True shape, Auxiliary view method, Development

1. INTRODUCTION

Many students majoring in product design are not good at understanding and using auxiliary view method. Especially for the students with hearing impairments, it is more difficult to teach it because they are said to have difficulty understanding abstract concept [2]. We must use sign language when we teach but it is difficult for us to tell the details on spatial relation fully with it.

Drawing for true shape is one of the methods made by using auxiliary view method which is also difficult to explain with sign language or the conversation by writing.

2. INVESTIGATION

We made a trial to five 3rd grade students (Table 1) who could not draw same kind of drawing six months before (Our university has 15 students in a department per a grade, and in that, 5 students major in product design). Their

numbers of right answer of MCT were 6-9 in 25 questions. It is equal to the result of junior high school students according to Tsutsumi's research [3].

Table 1: Student classification

Student	Coming High School	Result of MCT
A	High school of technology	8/25
B	Liberal high school	9/25
C		6/25
D		7/25
E	High School for deaf	6/25

In case of student B, majoring product design was the wish since the beginning. In case of students C, D, majoring graphic design was the wish since the beginning. Student E had been learning in deaf school from elementally to high school, and wished graphic design.

2.1 Problem for Construction

The problem for construction was as follows.

- 1. Making of top and front-view by using given perspective and right-side-view (Figure 1).
- 2. Making of developments of the planes which are not shown as true shapes on the drawing with the auxiliary view method in order to make paper-model (Figure 4).

This solid shape was designed as a dust-box set in the “portable vacuum cleaner”, which was the design theme of “basic design subject: Product Design C”.

2.2 Point of Instruction

The points of instruction were as flows.

- Let them find the plane where true shapes are not shown. And let them think and understand how those planes’ locations are in the spatial relation.
- Let them think how the true shapes should be made? (Let them notice the auxiliary view method out of many drawing methods they learned already.)
- Let them know if they cannot find the way to make true shape, they will not make paper-model.
- Let them make 3D-model with CAD, and make sure the true shape from the direction where it can be seen as the front view (Figure 2).

In the 3D-model, an arrow is added, which shows the direction to see the slant surface as true shape. Students can see it from various angles by themselves moving the mouse.

We intended to make the situation that if our students could not make this development, they would not be able to make the model completely. By this process, we expected that they found out the way to solve this problem by themselves not only as one of the drawing for the problems for construction but as the way to complete model correctly for their following

design works.

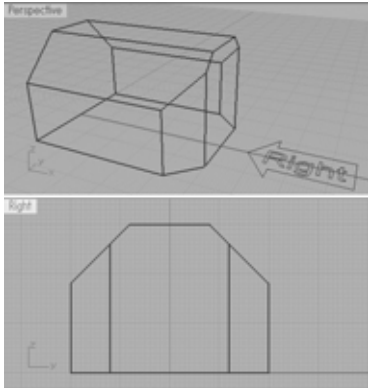


Figure 1: Figure shown to students

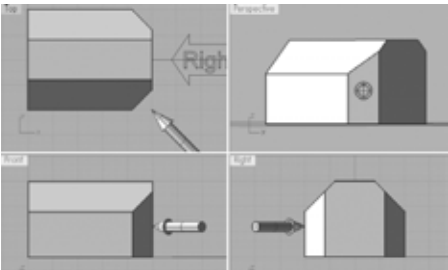


Figure 2: Additional figure

3. RESULT OF INVESTIGATION

As a result, 2 out of 5 students could draw correctly. Figure 3 shows one of those drawings. Though there were some poor parts and took much time, their drawings were according to the proper concept of its drawing method basically. Table 2 shows the students’ result of Drawing and MCT which were done before this drawing.

Table 2: Situation of achievement

Student	Result of Drawing
A	Correct
B	Correct
C	Incorrect
D	Incorrect
E	No finish

3.1 Situation of achievement

The situation of achievement of each student was as follows.

- In case of student A and B, most of the process was alright, but the place of auxiliary line was not correct. (It seemed not influence the results.)
- Student C and D could not understand the meaning of true shape and where those were, but they could remember the method itself finally.
- Student E could not notice the meaning of true shape, where those are, and also the method itself.

3.2 Process to the completion

The students' process to the understanding and completion of drawing was as follows (based on the observation and interview to the students who draw correctly after finish drawing).

1. They noticed there were two planes where true shapes were not shown in this drawing, but they seemed not understand what to do for a while.
2. After considering in about 15 or 20 minutes, they noticed from which direction the true shape could be seen, and remembered the auxiliary view method, which they learned a year before, could be applied.
3. They recognized the direction to see the plane as true shape, and recognized the length of lines which are set parallel to the direction were equal to the neighboring plane (without applying the method itself automatically).

In this process, we asked them when the slant surface was seen as true shape on the screen of CAD, and at the same time, how the arrow in it was seen.

Then, they seemed to notice that true shape was shown when the arrow was seen as a point. And next, they noticed the width and height of the figure (true shape) in that situation.

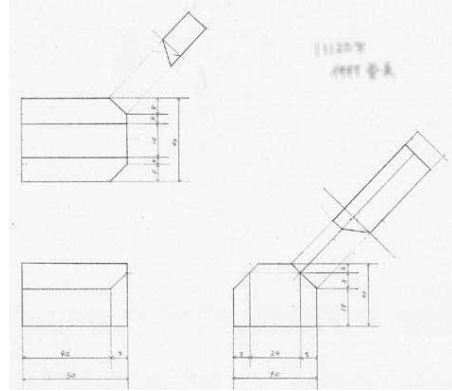


Figure 3: Example of correct drawing

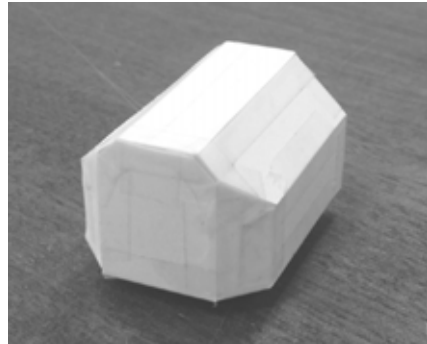


Figure 4: Completion of paper model

4. ANALYSIS OF RESULT

The key points of this result thought to be as follows.

- They thought by themselves not for solving problem but for making correct model. This point of view is thought to make them consider the way to draw in order to complete modeling.
- They noticed that drawing for the shape seen in the front might be drawn on the neighboring space which is divided with it. (Figure 5)

They also noticed that in the both plane neighboring "side-view (Figure 5)" the

length of depth from the top of the figure is same.

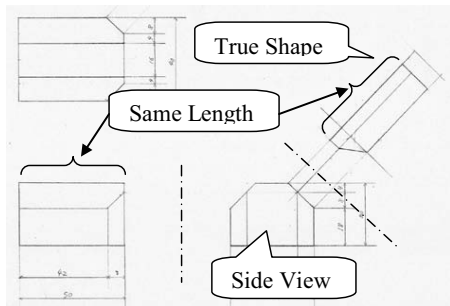


Figure 5: Relation of neighboring plane

For this, showing by CAD thought to help them, who are not good at understanding the explanatory notes, understand the spatial relation directly independent on the words.

As a result, we thought to be able to make our students notice (or remember) the similarity to the drawing process they learned past. And those who could use the auxiliary view method correctly seemed to understand it realistically and practically, which, we think, is the reproducible understanding.

5. CONCLUSIONS AND PROSPECTS

By this investigation, we confirmed that it is not enough to teach the drawing method as only a drawing process. It would be more effective to teach students a drawing method as a useful method if they know it and can use it. In order to understand it, it is necessary for them to have experience that what they learned is useful (for example, useful to complete a modeling).

Nevertheless, three of the remainder could not finish drawing. Their ability to imagine three-dimensional situation from given projected figures is found to be extremely poor according to the close inspection of their result of MCT (Table 1). And the educational environment which they belonged to before entering our university seems to have influence to

this result (Table 1).

For them, in the future, we must develop more useful method to notice the spatial relation.

REFERENCES

- [1] I. Honma. Problems in education of graphic science: auxiliary method. In *proceedings of the 2013 Asian Forum on Graphic Science (Dalian, China, August 9-11)*, pages 117-124. AFGS, 2013.
- [2] A. Yokkaich, S. Saito, and N. Tan. Vocabulary of children with hearing impairments: evaluation by item response analysis. In the Japanese journal of special education, pages 51. Japanese Association of Special Education, 1992.
- [3] E. Tsutsumi, T. Hongo, H. Yano, T. Yamamoto, and K. Suzuki. Investigation about spatial ability of the children and students: about the point of MCT. In *proceedings of JSGS autumn (Tokyo, Japan, December 15-16)*, pages 27-32. JSGS, 2012.

ABOUT THE AUTHORS

1. Iwao Honma: Professor at Faculty of Industrial Technology, Tsukuba University of Technology. Born in 1957. Graduated from the Department of Design, Chiba University. Began working at Honda R&D Co., Ltd. in 1982 as an exterior designer for passenger cars. Left Honda in 1988, then worked at some design studios before taking present post. Specialized studies are focused in product design and three-dimensional modeling education.

2. Kazunori Kuwabara: Teacher at Department of Information Design, Hokkaido High School for the Deaf. Born in 1972. Graduated from the Department of Education, Hirosaki University. Specialized studies are focused in education of basic modeling design. From 2000 to 2008, worked as a graphic designer and a teacher at the school for handicapped children.

AN EFFECTUAL OFFSET ALGORITHM FOR 3D CURVE IN ENGINEERING APPLICATIONS

Wei Liu, Chuipin Kong, Qiang Niu and Xionghui Zhou

National Engineering Research Centre of Die and Mold CAD, Shanghai Jiao Tong University

ABSTRACT: The offset of arbitrary curves is a typical issue in geometric modelling, which is widely applied in many engineering fields. But the offset of a rational curve usually cannot retain the same curve type as the origin one except for very few cases. Therefore the approximation method of curve offset is adopted for the uniform process of various types of curves. This paper presents a novel algorithm to dexterously discrete/sample the curve and to effectually control their precisions. Firstly sampling points are chosen according to such variables as offset distance and the distribution of control points to obtain initial offset. And then error of the resultant curve is calculated. If the error exceeds the pre-defined threshold, more points are sampling from where the error is beyond the scope of the local accuracy. These steps are repeated until the offset meets the requirements of precision. After extensive testing, compared with the traditional algorithm, our method can result in the same offset with less control points and lower fitting order. The algorithm has been successfully applied to the secondary development tools for CAD/CAM systems and improved the efficiency and accuracy of design.

Keywords: Offset, 3D curve, Discretization, Error control

1. INTRODUCTION

The offset of a curve is defined as another curve composed of a set of points having a given distance from the original one along the offset direction. This problem can be mathematically described as: $C^o(u) = C(u) + d \cdot N(u)$ where $C(u)$ is the original curve which can be 2D or 3D type, $C^o(u)$ is the offset one while d is the offset distance and $N(u)$ is the normal offset direction. Curve offset is a classical problem in the field of CAD & CAGD. In the past thirty years, researchers have been studying new algorithms and improving the algorithms because of the important use of curve offset in engineering applications such as art design, CNC machining tool path generation, path planning of robot, die and mold design, etc. In general, it is very difficult or even impossible to obtain the accurate solution of an offset-

ting curve except that the original ones are of handful types such as lines, arcs or hodographs, etc. The primary cause is that the normal can rarely be represented in the rational format. Moreover, most of the current researches are emphasized on the approximation algorithms for 2D cases. Corresponding 3D cases are more troublesome and hold several difficulties since their mathematical expressions are more complicated.

This status lies in the following factors. The offset direction of a curve usually doesn't share a fixed expression and it is difficult to forecast the self-intersection and overlap of the offset curve; the first-order discontinuity of the original curve may lead to the fracture problem; it's very difficult to control the precision of the offsetting curve. Since it is impossible for most curves to get a rational offset solution, an alternative way is to discretize them into several points, then the offset of each single point is computed, and finally the resultant offset

curves are reconstructed by interpolating with these offsetting points. Obviously, this technology is an approximate and compromise method which may bring certain deviation from ideal result.

To obtain a good solution of offset by the strategy of discretization, four key issues should be taken into serious consideration respectively. They are: error control, relation between offset distance and the sampling density, elimination of self-intersection, smoothness of offsetting curves.

2. RELATED WORK

Accurate offsets are limited for very few types of curve such as lines, arcs, pythagorean hodograph, and so on^[1-2]. Offset of most types of curves can only be obtained by approaching, which can be divided into 3 categories:

- I) Control Vertex Offset Method^[3-4]. Blomgren^[5] first proposed an implementation of curve offset under certain precision by offsetting control points of free curve. Tiller and Hanson^[6] improved Blomgren's algorithm by instead offsetting each edge of the control polygon of the curve in the normal direction, and generating the intersection of control polygon. The method is intuitive, shares the same expressions of approximating curve with the original one and holds an obvious geometrical meaning. However, the error control more difficult.
- II) Fitting and Interpolation Method^[7-9]. An ordered point set is created from the subdivision of original curve, each point is offset, and finally an offset curve is formed by interpolating these points. This method is simple and easy to implement, many scholars have deeply investigated it, of which Piegl and Tiller's algorithm^[10] is the most classic one. A NURBS curve is subdivided into several Bézier ones. For each Bézier curve, points of certain amount are sampled according to the maximum second derivative and the offset curve is fit from these points. Sun^[11] improved this method by de-

tecting the curvature radius in advance to avoid local self-intersections.

- III) Envelopment Method. Lee^[12] proposed a method to approach a circle using two order Bézier curves, and then compute the enveloping line swept by isometric approaching curve along the basis one. Based on Lee's algorithm, Zhao^[13] improved it on error measurement. Though offset curves obtained by envelopment method remain rational parametric form, their degrees are high, which brings a difficulty for the uniform representation in modelling systems.

To sum up, there are still some problems in the current algorithms of offset. Firstly, the error estimation method is relatively simple, underestimations may take place on many occasions; second, the control points are redundancy and the orders of offset curve are high, which is not conducive to the applications in engineering; Thirdly, the sampling does not take into account the offset distance, which perhaps lead to big errors.

3. OUR SOLUTION

Aiming at the problems described in Section 2, this paper puts forward the following solutions:

- I) Error function is adopted to fit the error of the resultant offset curve, which can help effectually analyse the distribution error and obtain the maximum error in an more accurate way;
- II) According to the error requirement, stepwise method is adopted to insert sampling points, so as to avoid redundant sampling point;
- III) The initial sampling density is related to offset distance to satisfy the different cases and reduce the complexity of the algorithm.

In this paper, we propose a novel unified 2D and 3D curve offset algorithm which intends to solve the problems by the following ideas: 1. Offsetting curve has the same precision by given the offset distance in the sampling formula, although its offset distance is different. 2. The self-intersection of the offset-

ting curve can be trimmed by checking the curvature or maximum curvature of the original curve at the sampled points for 2D curve. As for the 3D cases, they have to be projected onto a plane firstly and then check the curvature or maximum curvature of the projected one. This method can eliminate the self-intersection and overlap from the normal direction of projection plane. 3. Choosing the spherical Bézier curve to fix the fracture of the convex curve and cylindrical Bézier curve for the concave curve. 4. Fitting the offset curve error with a NURBS function which has the same node vector and weight with the original curve. With the help of the error function, it's easy to insert the points in the place where the offset error is out of allowed range.

3.1 Error Function

The traditional method of error estimation is to sample error on discrete points, the maximum value of which is regarded as the global error of offset. To accurately measure the error, too many points need be sampled. Obviously, its shortcoming lies in large computation and the result is smaller than practice. In our implementation, error distribution is fit as a function using the format of NURBS, in which nodes and corresponding weights remain the same as original curve. Once the error function is obtained, the regions where their values exceed the threshold and their nodes are easy to be found as shown in Figure 1. For those areas beyond threshold, the sampling density is

re-chosen in a higher level and the offset curve is then reconstructed. The current method is to simply insert a point into two adjacent ones and repeat the step until the whole curve meets the accuracy requirements.

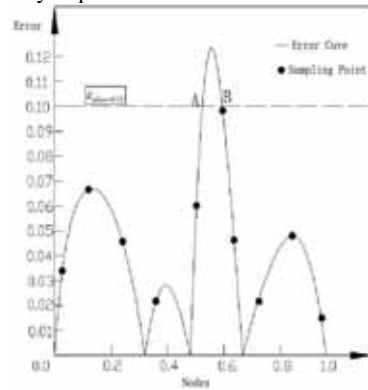


Figure 1: Error estimation of sampling points

3.2 The Relation between Offset Distance and Sampling Density

Usually the initial sampling density is decided according to Piegl's method, which didn't take offset distance into consideration. As shown in Figure 2, if the original curve is a straight line, the curvature of every point is 0, so the length of curve segments do not change after offset; if the original curve is convex along the offset direction, the distance between two subdivided points will be amplified; or else the distance will be shortened.

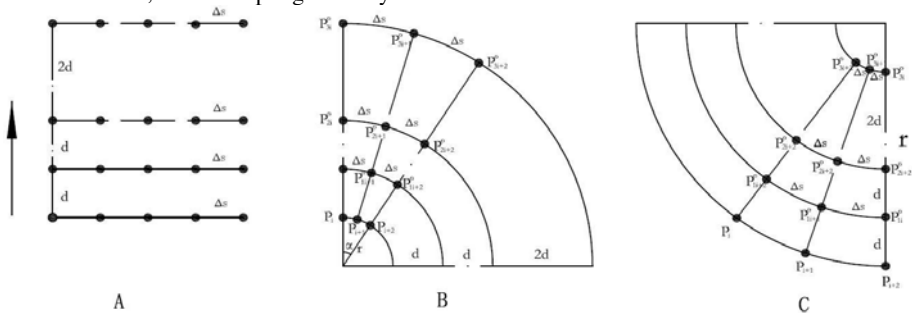


Figure 2: Three Examples of Offset A) line B) Convex C) Concave

The final offset curve is generated by interpolating the offset points. For the same

original curve, the error level is related to offset distance as shown in Figure 3.

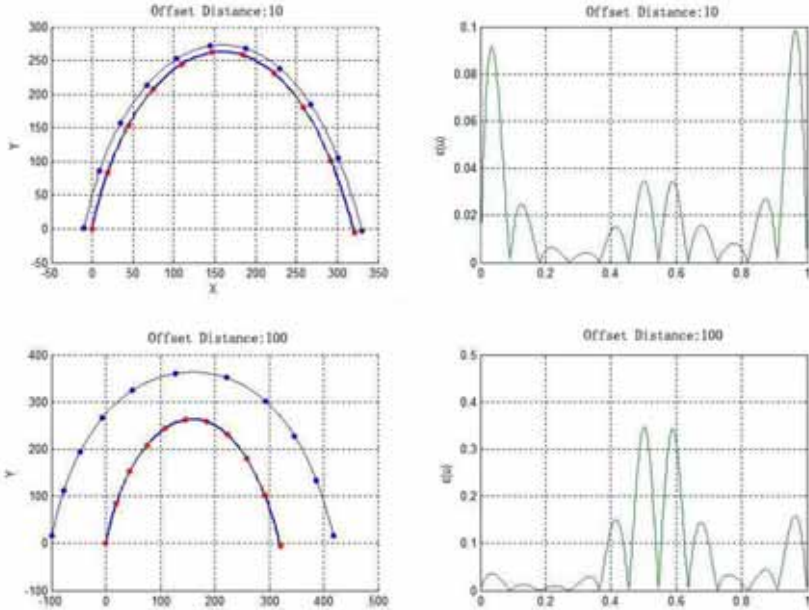


Figure 3: Error Distribution with Different Offset Distance

in the offset process.

For the same sampling point, Figure 3(top) illustrates the offset result and the error distribution with offset distance being 10, while Figure 3(bottom) illustrates that of distance being 100. It can be easily obverted that the maximum error of the former is less than 0.1 while that of the latter is greater than 0.3. Therefore there is a direct relationship between precision and offset distance, which should be considered as a factor for point sampling.

If the offset distance is smaller than its local curvature in an offset case, there will be a phenomenon of self-intersection, shown in Figure 4. An invalid loop together with a self-intersection point and two quasi self-intersection points will appear on such occasions. The loop is created since the curvature between point A and point B is too big. Therefore, a special measure should be taken to handle the segment between the two points



Figure 4: Self-Intersection

3.4 Our Offset Algorithm

On the consideration of above analysis, we present a novel offset algorithm. Firstly those segments where the curvature is greater than offset distance is removed. And then construct the initial offset curve based on a given number of sampling points which is determined according to offset distance and control points on the original curve. Error function is used to determine the number of points where the error exceeds the threshold. Since NURBS

curve is the de facto standard in CAGD, it is adopted as the example in the detailed analysis.

- I) Detect the curvature of the original curve with iterative method and record the corresponding regions $[u_k, u_{k+1}]$ and nodes $Q = \{Q_1, Q_2, \dots\}$;
- II) Sample points on the curve and join in $Q = \{Q_1, Q_2, \dots\}$, the initial number of sampling points is set to be $N = \left\lceil \frac{d}{r} \right\rceil h$, where d is offset distance, r is the maximal radius of curvature, and n is the control points. The points are uniformly sampled according to the pitch;
- III) Compute the offset point set $P = \{P_1, P_2, P_3, \dots\}$ for $Q = \{Q_1, Q_2, Q_3, \dots\}$ according to the actual offset definition;
- IV) Interpolate a NURBS curve with the same order as the original one according to the resultant set $P = \{P_1, P_2, P_3, \dots\}$ as the in-

tial offset curve, whose nodes are corresponding to the original one;

- V) Calculate the error distribution of the offset curve and form the error function, check if there are regions where their errors exceed the threshold, jump to VII if not;
- VI) If the error for region $[Q_i, Q_{i+1}]$ is greater than threshold, compute the point needed be inserted into $P = \{P_1, P_2, P_3, \dots\}$ according to analysis model for error;
- VII) Smooth the obtained curve.

4. EXPERIMENTS AND RESULTS

We compare our algorithm with Lee's, mainly focusing on the number of control points. The first example is a Bézier curve with the offset distance being 1.0. Figure 5 illustrates the result and Table.1 is the number of point in comparison with other methods.

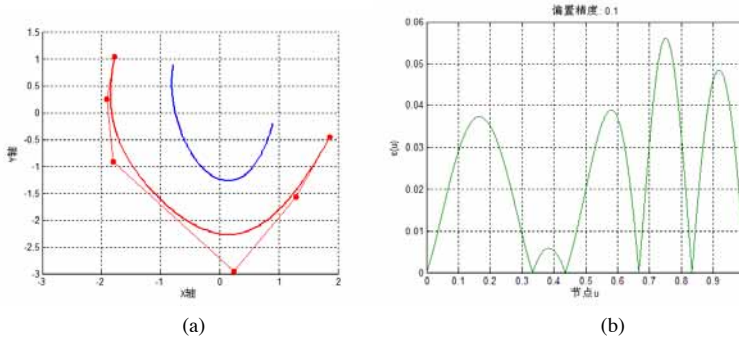


Figure 5: Bézier curve/its offset curve and error distribution of the offset curve with $\varepsilon=10^{-1}$

Table 1: Number of control points in comparison with other methods for the Bézier curve

ε	Cobb	Elb	Elb-2	Till	Lst	Lst-2	M-2	Our method
10^{-1}	10	11	13	10	7	10	22	6
10^{-2}	31	24	25	31	13	19	29	11
10^{-3}	94	74	97	97	19	31	43	18
10^{-4}	316	216	322	322	31	46	71	34
10^{-5}	865	974	769	886	50	88	127	48

Figure 6 shows an offset of a B-Spline curve and the distribution of errors on control points when we set the total error be $\epsilon=10^{-2}$. Table 2 shows the number of control points in comparison with other methods for the B-spline curve.

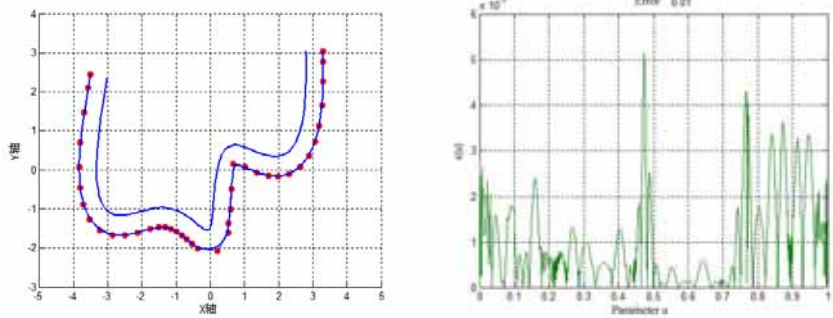


Figure 6: B-spline curve/ its offset curve and error distribution of the offset curve with $\epsilon=10^{-2}$

Table 2: Number of control points in comparison with other methods for the B- spline curve

ϵ	Cobb	Elb	Elb-2	Till	Lst	Lst-2	M-2	Our mehod
10^{-1}	28	19	22	25	16	31	78	17
10^{-2}	73	57	55	67	48	49	92	40
10^{-3}	208	174	190	202	84	84	120	67
10^{-4}	637	417	550	640	138	138	176	116
10^{-5}	1846	1357	1690	1918	240	240	302	197

This algorithm has been used in the design of checking fixture for automobile parts. In general, the detection surface of checking fixture flushes with detection surface of parts. Only extending the parts of the detected surface will appear some problems, such as surface branching and overlap as shown in Figure 7. Therefore,

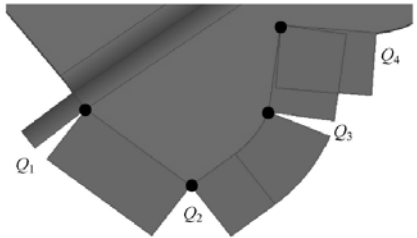
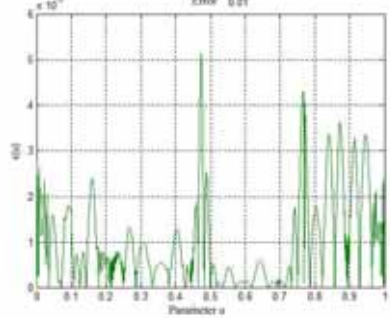


Figure 7: Branching and intersection/ overlapping after extending

From the table it can be observed that our method can obtain the same accuracy with less control points. Since we use less control points in computation, our method is within the class of higher efficiency.



this paper proposed to offset the parts' outline and then generate the detection surface with the help of the original curve and the offsetting curve(as shown in Figure 8). The algorithm is proved to be reliable and practicable both in geometrical theory and in engineering applications through several cases.

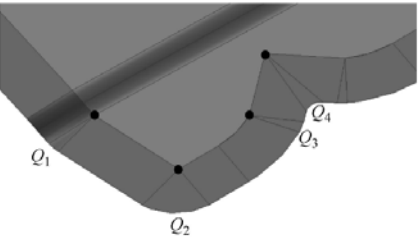


Figure 8: Detection surface generated by parts' outline and its offset curve

5. CONCLUSIONS AND FURTHER WORK

This paper proposed a new algorithm of offset by discretizing the original curve into several control point so as to convert the offset of curve into that of points. Error function is adopted to analyse the distribution of error, according to which the sampling points are accurately chosen. This algorithm can effectively reduce the number of iterations and quickly obtain the offset curve in high precision with less points. Finally, the algorithm is applied to the checking fixture for automobile parts, which effectively improved the efficiency and quality of design. This algorithm holds a widely application prospect in the CAD/CAM.

REFERENCES

- [1] Farouki R.T, Sakkalis T. Pythagorean Hodographs, IBM Journal of Research and Development, 34(5): 736-752, 1990
- [2] Zbynek S, Bohumir B, Miroslav L. Hermite Interpolation by Hypocycloids and Epicycloids with Rational Offsets, Computer Aided Geometric Design, 27(5): 405-417, 2010
- [3] Liu L.G, Wang G.J. Optimal Approximation to Curve Offset Based on Shifting Control Points. Journal of Software, 13(3): 398-403, 2002
- [4] Coquillart S. Computing Offsets of B-spline Curves. Computer Aided Design. 19(6): 305-309, 1987
- [5] Blomgren RM. B-Spline Curves, Class Notes, Boeing Document B-715-BB-WP-281/D-441.2, 1981
- [6] Tiller W, Hanson E.G. Offsets of Two-Dimensional Profiles. IEEE Computer Graphics & Application, 4(9): 36-46, 1984
- [7] Jiang P, Wang J. Rational Approximation of the offset Curves of Said-Bézier Curves. Journal of Computer Aided Design and Computer Graphics, 20(11): 1494-1499, 2008
- [8] Chen G.D, Cheng M, etc. Rational Approximation of Offset Curves by Parametric Speed Approximation. Chinese Journal of Computer, 25(9): 1001-1007, 2002
- [9] Lee E. Contour Offset Approach to Spiral Tool Path Generation with Constant Scallop Height. Computer Aided Design, 35(6): 511-518, 2003
- [10] Piegl L.A, Tiller W. Computing Offsets of NURBS Curves and Surfaces, Computer Aided Design, 31(2), 147-156, 1999
- [11] Sun Y.F, Nee A. Y. C, Lee K. S, Modifying Free-formed NURBS Curves and Surfaces for Offsetting without Local Self-intersection. Computer Aided Design, 28(8):1161-1169, 2004
- [12] Lee, I.K., Kim, M.S, Elber G. Planar Curve Offset based on Circle Approximation. Computer Aided Design, 28(8):617-630, 1996
- [13] Zhao H.Y, Wang G.J. Error Analysis of Re-parametrization based Approaches for Curve Offsetting. Computer Aided Design, 39(2):142-148, 1996

ABOUT THE AUTHORS

1. Wei Liu, born in 1979. Associate Professor of Shanghai Jiao Tong University, research interests include CAD/CG.
2. Chuipin Kong, born in 1988. Engineer of Shanghai Jiao Tong University, research interests include CAD, standardization of parts.
3. Qiang Niu, born in 1980. Assistant Professor of Shanghai Jiao Tong University, research interests include intelligent design of die and mold.
4. Xionghui Zhou, born in 1963. Professor of Shanghai Jiao Tong University, research interests include CAD/CG, intelligent design of die and mold, informatization of manufacturing industry.

ENERGETICAL STREAMS VISUALIZATION USING INTERACTIVE SANKEY DIAGRAMS

Daniele Ciminieri¹, Raphaël Ménard², Guillaume Meunier², Paolo Ciuccarelli¹

¹Politecnico di Milano, Italy ²Elioth, France

ABSTRACT: The topics of energy and environment are nowadays much discussed in the research context, as issues like climate change and global pollution become more and more pressing; such knowledge is also relevant for common people though, as it concerns the whole natural environment. Communicating the research is not easy, as the topic itself include the description of complex systems and processes. In this paper a novel, interactive method for representing energy fluxes is described, based on the well known Sankey diagrams. First, the notion of energy and its underlying conceptual model will be introduced. Every stream needed by life and society can be described as debit, credit and stock. These metabolic patterns are the foundation of the systemic model “the limits to growth” formalized in 1972 by the Club of Rome. All of those streams are linked together in space and in time and create a balance that will be discussed in the paper.

Sankey diagrams will then be introduced: invented by Riall Sankey to analyze the thermal efficiency of steam engines, they have since been applied to depict the energy and material balances of complex systems. This visual model is built upon a network graph made of nodes and weighted edges, and may be considered the easiest and most intuitive way to represent incoming and outgoing flows from arbitrary elements. The advancement in web technologies allows the development of interactive and modular visualizations that facilitate the comprehension of the scope and the complexity of the phenomenon, as well as their diffusion through a web application accessible by any modern browser.

“The Reforme energy simulator” is a web platform proposing a solution to the issues raised by this research, enabling the visual analysis and interaction with the energy fluxes taking place in a set of spaces. Users may explore a preset scenario focusing on the French and Parisian area ranging on 300 years, or submit their own data regarding an arbitrary region and get their respective visualization automatically generated.

Keywords: Energy, streams, diagram, Sankey, Data, visualization, web, simulator.

1. INTRODUCTION

More than two centuries ago, almost all of the energy consumed by humanity was provided by direct or derived forms of solar energy. In the first half of the 21st century, almost all of the world's energy consumption comes from our stock: coal, gas, oil, nuclear fissile.

Renewables are a very small minority in the

totality of the world used energy. The average conversion efficiency of solar energy irradiation remains exceedingly low: about 80W to 2100W consumed on average per person in the early 21st century. Yet on average about 170W / m² of energy lies on the earth's surface.

Most of our energetic needs are satisfied by our stocks. However, the problems caused by this approach (depletion of the non-renewable

energy sources, high CO₂ and other pollutants levels), push us towards a *resolarization* of our consumption model.

All the flows necessary for life and society can be described by a model composed of debit, credit and stock (see equation 1): water, energy, materials, population, economy, pollution etc. On the basis of this metabolic scheme, and in the context of a request of the Club of Rome, a team led by Donella and Dennis Meadows produced a world famous report in 1972, *The Limits to Growth* [7].

$$Offer - Demand = \Delta Stock \quad (1)$$

According to this relation, the inevitable end of energy stocks will impose an unavoidable equality between debit and credit. Therefore, a certain energy supply will forcibly be generated by:

- Solar energy and derived forms;
- Geothermal energy;
- Tidal energy.

On average and on a global scale, the deposit provided by the last two is negligible compared to terrestrial solar radiation.

1.1 The *Reforme* project

Reforme is an interdisciplinary research project participating to the *Ignis Mutat Res* programme organized by the French ministry of culture and communications. Its aim is to address the issue described in the previous section, analyze it and propose tools and novel approaches to expose its critical points and foster new urbanization patterns.

The project is structured in three main sections: *REpresentation*, aimed at representing in a simple and didactical way the flows of demand, offer and stock in space and time. The analysis is focused on the French territory, scaling down to a single zone of Paris. The final output of this section, which is also the main topic of this paper, is the realization of an interactive web platform which will help the users understand the energetic landscape of the

area taken into account, and let them insert information about their own territories in order to visualize them.

The second section, *FORMalization*, aims at providing a mathematical formal model of demand, offer and stock to precisely define the qualitative dimensions of energy.

Finally the third section, *MEtamorphosis*, focuses on the future transformations that could be enacted to cope with the need of balancing energy demand and offer, ranging from a global scale (by presenting possible future scenarios) down to a local one (by proposing urban solar infrastructures that can relocalize the energy basins).

2. THE DESIGN PROCESS

As discussed above, one of the aim of the project is to develop a digital, interactive visualization of the energetic model described. The design of such tool was carried on by the DensityDesign lab (which is part of the *Reforme* team).

Based on past experiences in developing interactive visualization applications for a general public [1, 6], the platform structure as well as the views and the interfaces were discussed with the domain experts in the team, in order to understand the main goals that should have been achieved and the best approaches to reach them.

The main issue in the design of the visualization was to communicate the complexity of the proposed model while preserving its completeness, and to stress the importance of the consequences deriving from its perpetuation.

In order to introduce non-expert users to the topic and let them fully comprehend the problem starting from the data, a communicative approach [5] was chosen, hence graphically stressing the most important insights and avoiding a purely analytical visualization of the information.

Also, by providing multiple levels of interactions (more specifically exploration and reconfiguration, see [10]), we encourage the user to engage with the application and discover him-

self new meanings and interesting patterns deriving from the representation of data.

2.1 The Sankey diagram

Talking specifically about the most suitable way to visualize the energetic model, different visual tools were taken into account. Many diagrams have been developed to describe architectures based on flows; in particular, the Sankey diagram is one of the most widely used due to its completeness and ease of reading. It shows not only the transfers involved but also the quantitative distribution of values in the transfers, instantly returning an intuitive overview of the system.

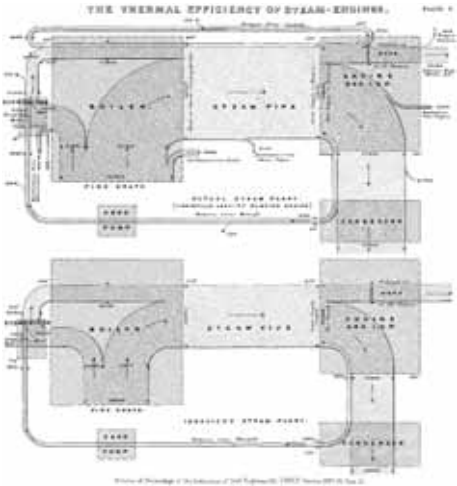


Figure 1 - Sankey's original 1898 diagram showing energy efficiency of a steam engine

A first prototype of this model can be seen in the "Napoleon's Russian campaign of 1812" map by Charles Joseph Minard, drawn in 1861. Such illustration shows the losses of the french troops during their march through the russian territory, and has been acknowledged by Edward Tufte as one of the best statistical graphics ever drawn [9].

The diagram itself takes its name from the irish engineer Riall Sankey, who used it to depict the energy flows of steam engines (see Figure 1). In more recent years, it was widely used to represent dynamic processes like transport flows [3] or trade exchanges [4].

A web tool to design interactive Sankey diagrams can be found in [8]: the application takes as input an XML file describing the data and renders the graph through a Java applet, enabling users to zoom, pan, select a particular flow to get additional details and animate the whole system.

We tried pushing further the technical implementation, using native web technologies. The visualization is drawn in SVG format, allowing its accessibility from computers as well as tablets and smartphones, while also maintaining an acceptable level of interaction.

2.2 Raw data

As discussed before, the data used to build the visualization is mainly focused on the french and parisien area, and ranging from 1750 to 2050 in five steps (1750, 1850, 1950, 2000, 2050).

In particular, for each time step, we take into account 5 spaces:

- Earth
- Europe
- France
- Paris
- Business district

Each space is contained into the previous one. For each step at each time, we have many information regarding its energetic equilibrium, i.e.:

- Population (absolute)
- Surface (in Km²)
- Stock (in TEP)
- Demand (in TEP)
- Offer (in TEP)
- Energy exchanges with the other spaces (in TEP)

THE PLATFORM

The platform is composed by two main modules:

- *The simulator*, a visualization of the model developed in the project focusing on the french area;
- *A user data driven section*, where information on a custom space may be inserted in order to be visualized.

Both modules rely on the same visualization model, built upon the data collected for the *Re-forme* project. In this main view the screen is divided into three main horizontal sections, described below.

2.3 Flows and exchanges

The first section is the Sankey diagram of the flows of energy exchanged between the spaces taken into account; the aim of this visualization is to let the user understand the generic architecture of the model, and to convey the role and importance of the renewable energies in the global energetic equilibrium.

The view is automatically generated for each year taken into account from the research project. Changes over time can be seen by selecting a specific year from the button set in the top left panel of the page.



Figure 2: Sankey diagram of the simulator

The fundamental building block of the view are the spaces, represented by horizontally aligned circles. Each space is a node of the Sankey diagram, hence having incoming and

outgoing energy flows. On the inside, each circle as two arcs (see figure 3).

The dark grey inner arc describes two values of the space stock: its amount S_p with respect to the world total for the current year (mapped to the angle of the arc), and its variation over time S_t (mapped to the height of the arc). Similar mappings are used for population in the outer green circles: P_p is the population percentage compared to the world total, while P_s represent its variation over time.

The flow O coming from the sun represents the offer, i.e. the solar energy actually exploited by the target space. Note that this is a minimal part of the overall solar radiation.

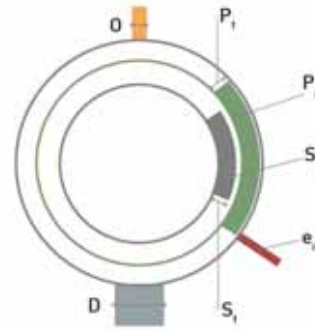


Figure 3: Detail of a territory in the Sankey diagram

The flow D going out from each space is the demand, i.e. the amount of energy needed by the territory. The sum of all the demand flows is the global consumption of energy. Finally, the red arcs e_i represent the energy net exchanges between spaces.

Note that all the energy data presented above (demand, offer, stock and exchanges) can be shown both as absolute values (measured in Joules) or as relative values by dividing them by the space population or surface amount.

By pointing the mouse or tapping on a specific arc or flow of the visualization, the user can see its value on a popup. Clicking or tapping on a space will instead cause it to merge

with its leftmost sibling (e.g. clicking on the Parc Des Ports space will cause it to merge with the Paris space). If a space is already grouped, clicking on it will cause it to expand and reveal the previously merged spaces.

The buttons *merge all* and *expand all* allow the user to see the union of spaces as a single territory (so that he can see the diagram relative to the entire world), and to reset the view to its default, i.e. with every single space shown.

2.4 Ins and outs

The second section is focused on the single spaces, showing the details of their incoming and outgoing energy amounts. Each space is represented by a skyline image (see figure 4); once again the D and O vectors are shown, indicating the demand and offer of the territory (or their density). Exchanges are also shown, but in this section they are grouped in two main arrows: E_{in} (incoming exchanges) and E_{out} (outgoing exchanges).

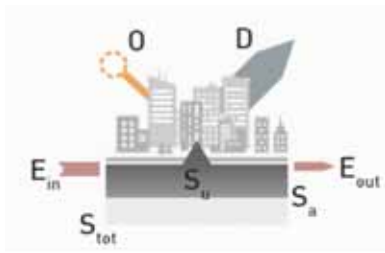


Figure 4: inputs and outputs of a space

Under the space the user can find visual information about the territory stock. The light grey area S_{tot} represents the maximum value of stock over time, while S_a is its actual value. Finally, S_u is the stock amount used to fulfill the demand, i.e. $S_u = O - D$.

Interaction is similar to the previous section: by hovering or tapping on an arrow or area, the user may read its value as a number in a popup; clicking on a territory causes it to merge with its leftmost sibling, and ungrouping it also pos-

sible as described before.

2.5 Trends

The aim of this section is to guide the user in finding connections between the trends of the values shown in the previous sections over time. The correlation between population, solar yield and stock variation is clearly visible, as the lines of each variable are drawn on the same graph (see figure 5).

Note that the graph is purely qualitative and the scales of the variable values are not respected. The view acts only as a comparison between the trends, to understand how the variation of a variable affects the others.

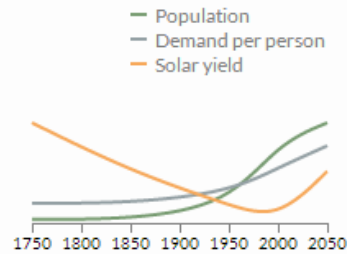


Figure 5: Line graph of Population, demand and solar yield over time

3. USER DATA DRIVEN VISUALIZATION

As discussed before, a second module of the platform lets the user build his own visualization by inserting data of a custom space. This feature is important not only to get a more complete understanding of the model, but to see how it fits for an environment that the user already knows. By looking at the visual output resulting from personal data, one can see new perspectives and find hidden issues of his territory (how much stock energy is needed to fulfill the demand of my city? How much my country depends on foreign furniture? And so

on).

The data input procedure is provided by a set of interactive graphs (see figure 6). Each graph is bound to a variable of the model (offer, demand, stock, etc), and its evolution over time is determined by 5 main steps, corresponding to the years taken into account by the research (1750, 1850, 1950, 2000, 2050).

By moving the handle of a given variable in a given time, the user automatically changes its value in the model. Such change is reflected in real time in the visualization, which is a custom instance of the one presented in the previous section. Hence the final outcome gets built step by step, and may be modified at any time.

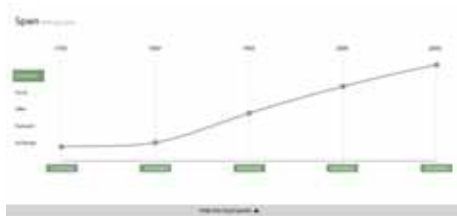


Figure 6: user data input view

Compared to filling a form, which is the standard way to insert data on the web, this graph-driven interface provides a more engaging and direct input method. This approach may be preferable when dealing with non expert users, as the speed and ease of use are relevant as much as precision and reliability.

4. TECHNICAL DETAILS

The *Reforme* simulator was developed as a web platform focused heavily on the front-end side. The need for complex visualizations and interactions was fulfilled by a heavy use of javascript: in particular the d3.js library [2] was adopted for the main views, while the angular.js library was chosen to handle the data underlying the views. In order to make the application usable from tablets and responsive to various screen sizes we used the Bootstrap CSS framework.

The back-end framework, Django, is a powerful yet easy to use application server written in Python. It was chosen for its seamless integration with the MySQL database and the possibility to quickly code user authentication and database administration pages.

5. CONCLUSIONS

In this paper, an interactive visualization for representing energy flows was presented. The application addresses a general public and aims at highlighting the importance of renewable energy sources in the current and future scenarios.

The data underlying the visualization is focused on the french and parisien zone, ranging from 1750 to 2050 in order to follow the evolution of the human footprint on environment.

In order to visually express the complexity of the energetic model, Sankey diagrams among other visual tools were used. Moreover, several levels of interaction have been included in order to foster the exploration of the application and the discovery of relevant patterns and findings.

In addition to the scenario taken into account as example, the user has the possibility to add information about an arbitrary territory and get the resulting visualization as output.

REFERENCES

- [1] Azzi, M., Caviglia, G., Ricci, D., Ciuccarelli, P., Bontempi, L., & Bonetti, E. Dust: A Visualization Tool Supporting Parents' School-Choice Evaluation Process. *Parsons Journal for Information Mapping*, 3(4). 2010.
- [2] Bostock, M., Ogievetsky, V., & Heer, J. D³ data-driven documents. *Visualization and Computer Graphics, IEEE Transactions on*, 17(12), 2301-2309. 2011
- [3] Cox D., Patterson R. *Visualization Study of the NSFNET*. <http://archive.ncsa.uiuc.edu/SCMS/DigLib/text/technology/Visualizaion-Study-NSFNET-Cox.html>

- [4] Krempel L., Plümper T. International Division of Labor and Global Economic Processes: An Analysis of the International Trade in Automobiles, *Journal of World-Systems Research*, Vol. V, 3, 487-498, 1999
- [5] Masud, L., Valsecchi, F., Ciuccarelli, P., Ricci, D., & Caviglia, G. From data to knowledge-visualizations as transformation processes within the data-information-knowledge continuum. In *Information Visualisation (IV), 2010 14th International Conference* (pp. 445-449). IEEE. 2010.
- [6] Mauri, M., Pini, A., Ciminieri, D., & Ciuccarelli, P. Weaving data, slicing views: a design approach to creating visual access for digital archival collections. In *Proceedings of the Biannual Conference of the Italian Chapter of SIGCHI* (p. 21). ACM. 2013.
- [7] Meadows D.H., Goldsmith E., Meadow P. *The limits to growth*. Universe books, New York, 1972.
- [8] Riehmann, P., Hanfler, M., & Froehlich, B. Interactive sankey diagrams. In *Information Visualization, 2005. INFOVIS 2005. IEEE Symposium on* (pp. 233-240). IEEE. 2005
- [9] Tufte, E. R., & Graves-Morris, P. R. *The visual display of quantitative information* (Vol. 2). Cheshire, CT: Graphics press. 1983
- [10] Yi, J. S., ah Kang, Y., Stasko, J. T., & Jacko, J. A. Toward a deeper understanding of the role of interaction in information visualization. *Visualization and Computer Graphics, IEEE Transactions on*, 13(6), 1224-1231. 2007.
3. Guillaume Meunier, project manager, Elioth.
4. Paolo Ciuccarelli, scientific coordinator at DensityDesign lab and associate professor, Design faculty, Politecnico di Milano.

ABOUT THE AUTHORS

1. Daniele Ciminieri, research assistant, DensityDesign lab, Design faculty, Politecnico di Milano.
2. Raphaël Ménard, director, Elioth.

ENUMERATION OF DELTAHEDRAL GRAPHS WITH UP TO 10 VERTICES

Naoya TSURUTA, Jun MITANI, Yoshihiro KANAMORI and Yukio FUKUI
University of Tsukuba, Japan

ABSTRACT: In this paper, we enumerate the polyhedral graphs that are realizable as deltahedra with up to ten vertices. We call these “deltahedral graphs”. This result was achieved by an experimental approach that trying to construct deltahedra from each of the simple cubic polyhedral graphs. We also provide examples of the graphs that are not realizable as deltahedra. We show that the infinite families of such nonrealizable graphs can be obtained by solving the graph isomorphism problem.

Keywords: Deltahedron, Polyhedral Graph, Geometric Realization

1. INTRODUCTION

A deltahedron is a polyhedron whose faces are congruent equilateral triangles. Only eight of these are convex: those having 4, 6, 8, 10, 12, 14, 16, or 20 faces [5]. Coplanar faces sharing an edge are not allowed. The tetrahedron, octahedron, and icosahedron are the three deltahedra that are regular solids. If we permit nonconvex shapes, then the number of deltahedra is infinite, because we can compose larger deltahedra by attaching two smaller deltahedra.

There are several subclasses of deltahedra. Cundy listed 17 biform deltahedra, which have only two forms of vertices [4]. Olshevsky added another 11 biform deltahedra to Cundy’s list [9]. These lists did not permit intersecting faces, so these biform deltahedra are solids. Shephard presented 34 isohedral deltahedra, and McNeill added six further examples to Shephard’s list [8, 13]. Isohedral deltahedra are face-transitive and may include intersecting faces. Trigg defined spiral deltahedra as those constructed from strips of equilateral triangles [14].

Each of these classes of deltahedra have their own particular properties. Therefore, the configurations of the vertices are very limited. We decided to loosen the conditions and see what kinds of shapes are possible in the world of

deltahedra. In this study, we tried to construct deltahedra from each of the simple polyhedral graphs and counted the graphs that can be constructed as deltahedra, that is, the deltahedral graphs. It is hard to determine whether a graph will form a deltahedron by examining only its structure. Thus, we solve a geometric realization problem, which is the problem of determining whether a triangulation of an orientable surface can be realized geometrically in R^3 as a polyhedron without self-intersections [6].

We present an enumeration of all deltahedral graphs with up to ten vertices and provide examples of the constructed deltahedra. In our realization process, we generate an initial polyhedron with nonequilateral triangles and then deform the faces into equilateral triangles by a gradient method, because the graph does not provide the locations of the vertices. Olshevsky focused on the operations used to construct the deltahedra [9]. Augmentation is an operation that joins each appendage polyhedron to its own single-core face. These simple operations can be detected by solving a graph isomorphism problem. We show that this is useful for finding an infinite family of graphs that are nonrealizable as deltahedra. Note that our result of an enumeration is not theoretical. We provide a realiza-

tion process that constructs a deltahedron from a graph, but it does not necessarily guarantee the nonrealizability of a graph.

Our deltahedral realization problem is a particular case of a geometric realization problem. In general, Bokowski and Guedes de Oliveira [1] showed that there is a nonrealizable triangulation of the orientable surface of genus 6, and Schewe [12] showed that we can construct non-realizable triangulations for any number of vertices genus 5 or 6. However, for surfaces of genus $1 \leq g \leq 4$, the problem remains open. The conditions for deltahedral realization are stricter than those. Each face must be realized as an equilateral triangle, and it is necessary to calculate the geometric coordinates to determine whether there exist self-intersections or edges whose dihedral angle is equal to 180° . In this paper, we focus on a surface of genus 0. Previous studies of deltahedra mentioned above also focused on the genus-0 surface. Although a few deltahedra with $g > 0$ are known, we know of no published studies.

2. DELTAHEDRAL GRAPH

Polyhedral graphs are three-connected planar graphs. These graphs contain not only triangular faces, but polygonal faces which have more than four edges. A cubic polyhedral graph is a three-connected cubic planar graph and has only triangular faces. This graph is realized as a polyhedron whose faces are triangles, that is, a simple polyhedron. Deltahedra are a subclass of simple polyhedra because they are composed of equilateral triangles. Therefore, the graphs of deltahedra are a subclass of the graphs of simple polyhedra. The relation between them is shown in Figure 1.

Here we define a deltahedral graph as a graph which can be realized as a deltahedron. Although there are various kinds of deltahedra, we will include only deltahedron that do not have any self-intersections and do not have any edges for which the dihedral angle is 180° . For example, all the polyhedra in Figure 2 are com-

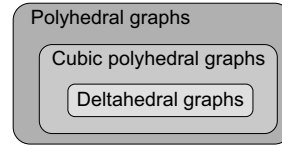


Figure 1: Relation between polyhedral graphs and deltahedral graphs.

posed of congruent equilateral triangles. However, the lower left one $(9_{24}, N)$ has intersecting faces, and the lower right one $(8_{12}, N)$ has coplanar and connecting faces. In this paper, we will not consider deltahedron like these lower ones but only those like the topmost example $(9_{17}, D)$ in Figure 2. The code below each figure is composed of two numbers and a character. The numbers represent the number of vertices and the index of the graph. The index follows the order of an existing graph generation algorithm [3], and they are classified into one of two categories: ‘D’ for deltahedral graphs, and ‘N’ for nondeltahedral graphs. For example, $(6_1, D)$ is the first six-vertex deltahedral graph that is generated by that algorithm. The important thing is that more than one deltahedra may be obtained from a single graph. Figure 3 shows a deltahedral graph that has one convex form and two nonconvex forms. If a graph has at least one deltahedron, we say it is deltahedral.

3. APPROACH

We enumerate the deltahedral graphs which are combinatorially different. The class of deltahedra is a subset of the class of simple polyhedra whose faces are triangles. Therefore, the number of three-connected cubic polyhedral graphs is an upper bound on the number of deltahedral graphs. Graph enumeration has been widely discussed, and there are many approaches to it. We used the planar graph generation program plantri [3] to obtain the three-connected cubic planar graphs.

We used two steps to realize the graphs. First, each graph was embedded without intersections

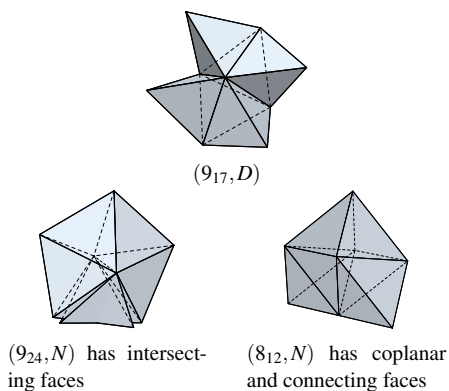


Figure 2: Polyhedra with equilateral triangles. The top polyhedron is a deltahedron and the others are not.

in the 2D plane with straight line edges. Then, we used graph lifting [11] and an iterative deformation process to attempt to construct deltahedra from the graph. A constructed polyhedron may have coplanar neighboring faces or may not even be a solid. The graph was considered to be a deltahedral graph only if the constructed geometry satisfied the conditions for a deltahedron.

As mentioned above, nonconvex deltahedra form an infinite class, and so in order to enumerate them, it is necessary to limit the allowable number of vertices. We chose this limit to be ten. As an upper bound, there are 306 polyhedral graphs which have ten or less vertices. We decided that is enough for the first step of this enumeration problem.

3.1 Graph Embedding

Several methods have been proposed for embedding planar graphs. We used Plestenjak's algorithm, which is based on a spring model [10]. The size of the graph is small enough that it is practical to calculate it. This algorithm chooses a base face and fixes the positions of its vertices in the 2D plane. The remaining vertices are placed inside the base face. We choose the base face randomly and place it so that it forms

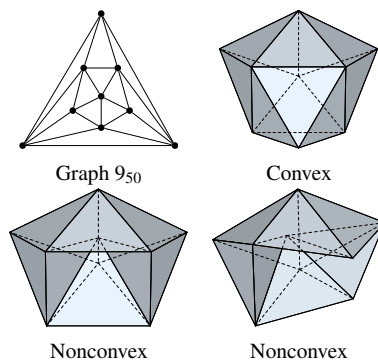


Figure 3: One convex shape and two nonconvex shapes form the graph 9_{50}

an equilateral triangle with edges of unit length. The algorithm calculates the periphericity p_v of each vertex when placing the inner vertices. Periphericity is a kind of centrality and indicates the distance from the outer polygon. The p_v of the outer triangle is 0, and the p_v of the vertices adjacent to them is 1. The p_v increases as going toward the inside. These periphericities are used when generating the initial polyhedron to be used in the iterative deformation. Figure 4(a) is an example of an embedded graph. In this step, the base face is chosen randomly. The results of following steps are different, depending on this initial choice.

3.2 Realization of Deltahedra

First, we generate a polyhedron with nonequilateral triangles from an embedded graph. Although this polyhedron will be deformed to a deltahedron, it should be close to a deltahedron. We obtained the heights h_v for the vertices corresponding to each p_v by using the following formula:

$$h_v = Cp_v$$

where C is a constant value that defines the height of the pyramid-like model shown in Figure 4(b). The base face was then shrunk to reduce the differences in the lengths of the edges

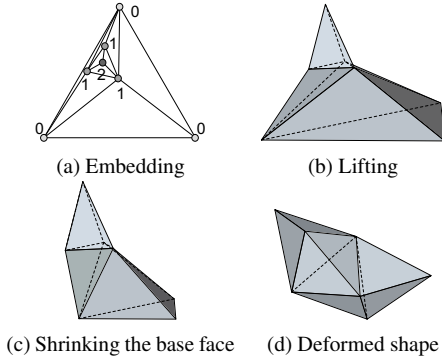


Figure 4: Realization process.

(Figure 4(c)).

Finally, we transformed the generated polyhedron to form a deltahedron by using numerical optimization. We defined a penalty function F which measures the difference in length of the edges, and we then minimized it:

$$F := \sum_i (L(E_i) - 1.0)^2$$

where $L(E_i)$ is the length of E_i . We used the Levenberg-Marquardt algorithm for the iterations and the Gauss-Seidel method for solving the linear equations within each iteration. The resulting polyhedron may include intersecting faces. When this was the case, we tried a different embedding or manually reconfigured the positions of the vertices.

Convergence of this method is not guaranteed. However, there exists a vertex configuration that makes all the lengths of the edges equal when self intersection is not considered. The proof is as follows. There are only three operations for generating all triangulations[2] including:

- a) Adding a vertex of degree 3
- b) Removing an edge and adding a vertex of degree 4
- c) Removing two edges and adding a vertex of degree 5

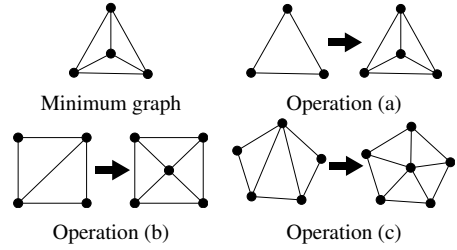


Figure 5: Operations for generation of simple triangulations.

The minimum four-vertex graph is realized as a tetrahedron. The graph and the operations are shown in Figure 5. When these operations are applied to it, the resultant polyhedron can be fit within an inside region of the tetrahedron because our iterative deformation algorithm does not care about intersecting faces. Operation (a) adds a vertex that separates a face into three faces. The additional faces can be realized as an excavation of a tetrahedron. The additional faces generated by operation (b) will be realized as overlapping coplanar faces resulting in turning over other faces. Operation (c) also causes the turning over of the whole shape, although the shape will still fit within the tetrahedron.

4. RESULTS

Here we present the results of the enumeration and reconstructions. Table 1 shows the numbers of deltahedral and nondeltahedral graphs. We can see that more than half of the graphs are deltahedral graphs. The percentage of nondeltahedral graphs is seen to gradually decrease. To confirm this trend, it may be necessary to investigate larger graphs. Figures 6 and 7 show the constructed polyhedra with seven and eight vertices, respectively. Note that each figure is one of various possible polyhedra. We did not enumerate all the possible realization shapes for each graph. Fortunately, if we do not allow faces to intersect, the variations are small in graphs with ten or fewer vertices. It is easy to manually identify whether a graph has different shapes. For

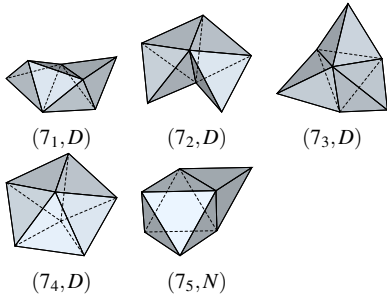


Figure 6: Constructed polyhedra with $V = 7$.

difficult cases, we manually generated *good* initial polyhedra and performed an iterative deformation.

The lengths of all the edges of the constructed polyhedra are very close to 1. Although there is an error caused by the numerical calculations, the maximum difference between the mean edge length and each edge was under 10^{-5} . In most cases, the iteration process converged in a few seconds on a PC with 2.9 GHz Intel Core i7 CPU. The time differed depending on the initial polyhedron produced by the graph lifting.

5. FINDING THE INFINITE FAMILY OF NONDELTAHEDRAL GRAPHS

Some graph structures cause face intersections. For example, in Figure 7, the shape is completely flat in the case of graph 87. Similar shapes appear in larger graphs. We can find the graphs that have the particular nonrealizable structures by comparing graphs.

Figure 8 shows graphs which contain the same partial structures and their realized polyhedra. As shown in Figure 8, when we form a larger deltahedra by connecting two smaller deltahedra along a single face, the original shapes do not change and the graph of the appended deltahedron can be embedded inside a connecting face. This operation can be detected by solving the subgraph isomorphism problem. Hence, we can obtain a family of nondeltahedral graphs from one nondeltahedral graph by solving it, without

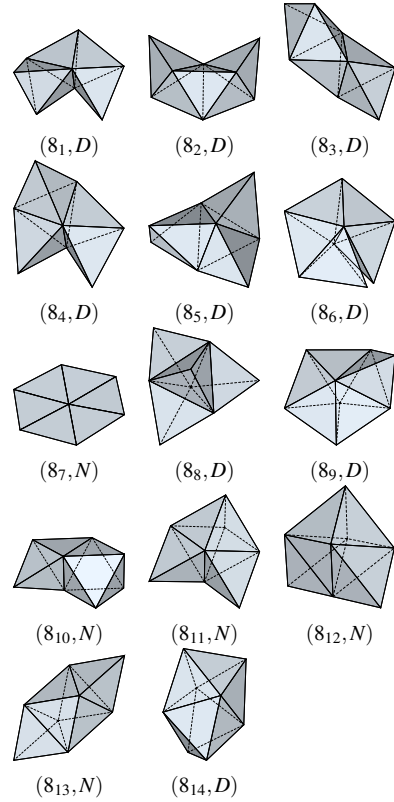


Figure 7: Constructed polyhedra with $V = 8$.

the need to realize the polyhedra.

Figure 9 shows an example of a nondeltahedral family. We used graph 87 as a seed and obtained ten nondeltahedral graphs with nine or ten vertices. We used a simple backtracking algorithm for the subgraph isomorphism [15]. The computation times were 140 ms and 2500 ms for nine and ten vertices, respectively.

This subgraph isomorphism only detects the connection of two deltahedra that involves a single face. Figure 10 is a comparison of a connection that involves only one face and with one that involves multiple faces. These shapes look similar, but the isomorphism of the subgraphs only

Table 1: Number of deltahedral graphs.

Vertices	4	5	6	7	8	9	10
Graphs	1	1	2	5	14	50	233
Deltahedral graphs	1	1	2	4	9	36	154
Nondeltahedral graphs	0	0	0	1	5	14	79

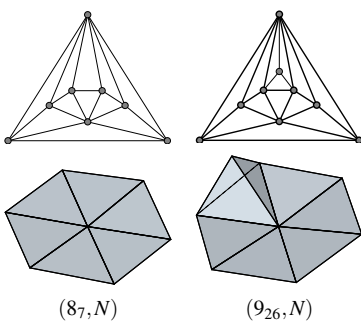


Figure 8: Graphs with and without the appendage tetrahedra, and the associated polyhedra.

detects the deformation from left to center. The right one is a seed of another nondeltahedral family. From center to right, an octahedron is attached along two faces. In this case, the original shapes do not change, but in general, attaching a deltahedron along multiple faces causes deformations in the original shapes.

The graph 7_5 can be realized as a polyhedron that does not contain self-intersections. Such nondeltahedral graphs that are realizable as polyhedra cannot be used as seeds of the subgraph isomorphism problem. A larger graph may be realizable as a deltahedron because the attachment changes the dihedral angles of the edges around the core face.

6. CONCLUSIONS

We have described a method for and the result of an enumeration of deltahedral graphs which have ten or less vertices. Not all simple cubic polyhedral graphs can be realized as deltahedra, due to self-intersections or dihedral angles

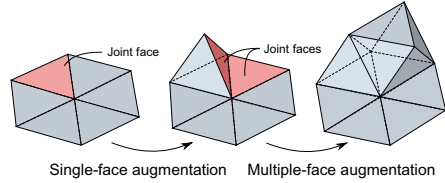


Figure 10: Connection with single faces and with multiple faces.

of 180° . We have also shown that the infinite families of nondeltahedral graphs are obtained by solving the subgraph isomorphism problem. This eliminates the nondeltahedral graphs from the set of cubic polyhedral graphs without the need for realization, and it may be useful for finding the deltahedral ones. Our future work is to improve the discrimination approach by employing this detection method.

The remaining problem is to determine how nonrealizability can be characterized. In order to do this, we need the vertex coordinates because of the dihedral angles. Our deltahedral realization problem is similar to the polyhedral realization problem. We hope that by combining our iterative deformation process with other realization or detection methods [7, 12], we will obtain a method that creates a robust realization of deltahedra.

It is also necessary to investigate graphs which have higher genus numbers. Do all triangulated surfaces with nonzero genus admit a deltahedral realization? Our realization process is not applicable for surfaces with nonzero genus; however, an iterative deformation may be useful. It will be an interesting challenge to find the smallest deltahedron with $g > 0$, such as a toroidal deltahedron.

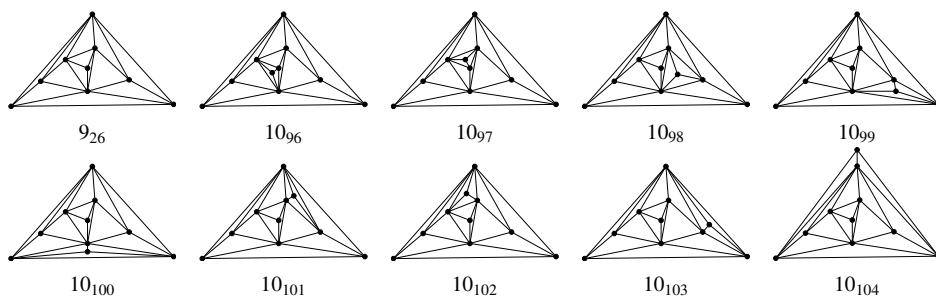


Figure 9: Nondeltahedral family of graph 8_7 up to ten vertices.

REFERENCES

- [1] J. Bokowski and A. Guedes de Oliveira. On the generation of oriented matroids. *Discrete & Computational Geometry*, 24(2-3): 197–208, 2000.
- [2] R. Bowen and S. Fisk. Generation of triangulations of the sphere. *Mathematics of Computation*, 21(98): pp. 250–252, 1967.
- [3] G. Brinkmann and B. D. McKay. Fast generation of planar graphs. *MATCH Communications in Mathematical and in Computer Chemistry*, 58: 323–357, 2007.
- [4] H. M. Cundy. Deltahedra. *Mathematical Gazette*, 36: 263–266, 1952.
- [5] H. Freudenthal and B. L. van der Waerden. On an assertion of euclid. *Simon Stevin*, 25: 115–121, 1947.
- [6] B. Grünbaum. *Convex Polytopes*, volume 221 of *Graduate Texts in Mathematics*. Springer New York, 2 edition, 2003.
- [7] F. H. Lutz. Enumeration and random realization of triangulated surfaces. In A. I. Bobenko, J. M. Sullivan, P. Schröder, and G. M. Ziegler, editors, *Discrete Differential Geometry*, volume 38 of *Oberwolfach Seminars*, pages 235–253. Birkhäuser Basel, 2008.
- [8] J. McNeill. Isohedral deltahedra. http://www.orchidpalms.com/polyhedra/rhombic/isohedral_deltahedra.htm.
- [9] G. Olshevsky. Breaking cundy’s deltahedra record. [http://www.interocitors.com/polyhedra/papers/Polytopics_28 - Breaking Cundy’s Deltahedra Record.pdf](http://www.interocitors.com/polyhedra/papers/Polytopics_28_-_Breaking_Cundy's_Deltahedra_Record.pdf).
- [10] B. Plestenjak. An algorithm for drawing planar graphs. *Softw. Pract. Exper.*, 29(11): 973–984, September 1999.
- [11] G. Rote. Realizing planar graphs as convex polytopes. In *Proceedings of the 19th International Conference on Graph Drawing*, GD’11, pages 238–241. Springer-Verlag, Berlin, Heidelberg, 2012.
- [12] L. Schewe. Nonrealizable minimal vertex triangulations of surfaces: Showing non-realizability using oriented matroids and satisfiability solvers. *Discrete Comput. Geom.*, 43(2): 289–302, January 2010.
- [13] G. Shephard. Isohedral deltahedra. *Periodica Mathematica Hungarica*, 39(1-3): 83–106, 2000.
- [14] C. W. Trigg. An infinite class of deltahedra. *Mathematics Magazine*, 51(1): pp. 55–57, 1978.

- [15] J. R. Ullmann. An algorithm for subgraph isomorphism. *J. ACM*, 23(1): 31–42, January 1976.

ABOUT THE AUTHORS

1. Naoya Tsuruta, Ph.D. student at Department of Computer Science, University of Tsukuba.
2. Jun Mitani, Associate Professor at Department of Computer Science, University of Tsukuba.
3. Yoshihiro Kanamori, Assistant Professor at Department of Computer Science, University of Tsukuba.
4. Yukio Fukui, Professor at Department of Computer Science, University of Tsukuba.

ERROR ANALYSIS OF PROBLEMS ON SURFACE AREA OF AN OBJECT PRESENTED IN A THREE-VIEW ORTHOGRAPHIC DRAWING AS A PART OF A MATHEMATICS TEST FORM

Kumiko SHIINA, Sayaka ARAI, Tatsuo OTSU

The National Center for University Entrance Examinations, Japan

ABSTRACT: Mathematics problems used for a university entrance examination in China were adapted into Japanese and administered to Japanese freshmen together with subject tests in the National Center Test, which was developed for university entrance examinations in Japan with reference to Japanese curriculum. Our interest is in test properties of the problems about an object presented in a three-view orthographic drawing for Japanese freshmen, most of whom had not studied three-view orthographic drawings. Thus a brief explanation of three-view orthographic drawings was added for subjects. It was revealed that the two problems contributed to discriminating between subjects with high and low mathematics scores. The subjects with low scores tended to choose wrong answers due to misunderstanding a triangle in each orthographic view as a true-size view. In addition, difficulties in understanding the true size of the face presented as an edge view in each orthographic view were suggested for the problem in the test form for science courses. Interestingly, a significant association was found between performance in the problems on three-view orthographic drawings, which are not taught in the mathematics curriculum in Japan, and the test scores for mathematics developed with reference to the Japanese curriculum.

Keywords: Graphics education, three-view orthographic drawing, error analysis

1. INTRODUCTION

Nation-wide examinations play an important role in selecting university entrance applicants in East Asian countries such as Japan, China, and South Korea. Thus the contents and statistical properties of nation-wide examinations in neighboring countries are of interest, although pre-university curriculums in these countries differ slightly.

Mathematics test forms used for a university entrance examination in China were adapted into Japanese [2] and administered to Japanese freshmen together with mathematics tests in the National Center Test, which was developed for university entrance examinations in Japan with reference to the Japanese curriculum [1]. Despite the difference in the constitution of test contents and the test items' response format, the correlation coefficients between scores in

mathematics test forms of both countries were quite high [1].

Among the problems of Chinese origin used in the investigation, this paper focuses on problems including an object presented in a three-view orthographic drawing. Unlike in China, three-view orthographic drawings are not taught in the pre-university mathematics curriculum in Japan. Our interest is in test properties of the problems about an object presented in a three-view orthographic drawing for Japanese freshmen, most of whom had not studied three-view orthographic drawings. Errors made by such examinees are also of interest.

2. INVESTIGATION

2.1 Research design

Mathematics problems used for the College

Entrance Examination (CEE) in the Beijing area in 2011 were adapted into Japanese [2] and administered to 348 freshmen at five national universities in Japan in January 2013 [1]. In addition, the subject test of the National Center Test (NCT), which was developed for university entrance examinations in Japan with reference to Japanese curriculum, was also given.

There are two kinds of mathematics test forms used for CEE in the Beijing area: one is developed for applicants for science courses and the other for applicants for liberal arts courses. The solving time limit for both original form is 120 minutes, and the full score for each is 150. Both forms contain multiple-choice items and constructed-response items.

Due to time constraints, each original form was divided into two forms with similar composition of contents and items' response format. One form, which was originally developed for applicants for science courses, was divided into R1 and R2 forms, and the other form for applicants for liberal arts courses was divided into R3 and R4 forms. One of the four forms was given to freshmen attending corresponding faculties in Japan. The solving time limit for each form was shortened to 70 minutes, and the full score for each divided form was 75.

2.2 Problems

Two of the four forms included a multiple-choice problem about the surface area of an object presented in a three-view orthographic drawing. Figures 1 and 2 show the problems focused on in this paper. Figure 1 shows the problem I-(4) in R1 form, given to freshmen in science courses. Hereinafter, this problem will be referred to as problem-S. In problem-S, examinees were required to choose the value of the area of the largest face of a given tetrahedron. Figure 2 shows the problem I-(3) in R3 form, given to freshmen in liberal arts courses. Hereinafter, this problem will be referred to as problem-L. In problem-L, examinees had to choose the value of the surface area of a given quadrangular pyramid. In both problems, the

examinees had to understand the shape of the object in order to choose the right answer.

In the original Chinese forms, a three-view orthographic drawing was presented using first-angle projection, and no explanation of the drawings was added. Unlike in China, three-view orthographic drawings are not taught in the pre-university mathematics curriculum in Japan. Instead, three-view drawings using third-angle projection are taught in technology and home economics in some junior high schools. Thus, three-view orthographic drawings using first-angle projection were changed into three-view drawings using third-angle projection, and a brief explanation of how to read three-view drawings was added for examinees who had not studied them.

Problem I-(4)

The following figure shows an orthographic drawing of a tetrahedron. An orthographic drawing consists of a front view, a top view and a right-side view, aligned horizontally and vertically; each of which shows a projection of the object from the far distance of front, top and right-side, respectively. Choose the value of the area of the largest face of the given tetrahedron.

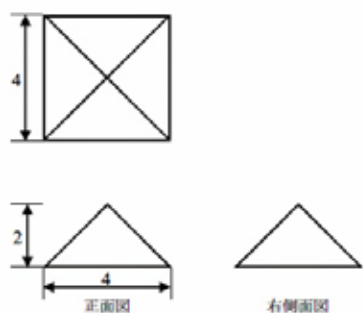
The figure shows three orthographic views of a tetrahedron. The top view is a triangle. The front view is a triangle with a vertical height of 4 and a horizontal base of 4. The right-side view is a triangle with a horizontal base of 3. The views are aligned horizontally and vertically.

①8 ② $6\sqrt{2}$ ③10 ④ $8\sqrt{2}$

Figure 1: Problem I-(4) in R1 form (Problem-S)

Problem I-(3)

The following figure shows an orthographic drawing of a quadrangular pyramid. An orthographic drawing consists of a front view, a top view and a right-side view, aligned horizontally and vertically; each of which shows a projection of the object from the far distance of front, top and right-side, respectively. Choose the value of the surface area of the given quadrangular pyramid.



①32 ② $16+16\sqrt{2}$ ③48 ④ $16+32\sqrt{2}$

Figure 2: Problem I-(3) in R3 form
(Problem-L)

3. RESULTS AND DISCUSSION

Analysis was conducted using data from 103 and 73 freshmen who were assigned the R1 and R3 forms, respectively.

3.1 Error analysis of problem-S

Table 1 lists the frequency and proportion of subjects choosing each alternative for problem-S. Figure 2 shows quantile item response charts for problem-S. The subjects are divided into four groups (Q1-Q4) with nearly equal numbers in descending order of total score for R1 form. The proportion of subjects choosing each alternative for each group is plotted. The proportion for correct alternative (Alternative 3) is plotted with a thick line. A quantile item re-

sponse chart indicates discrimination power of alternatives in separating the high and low groups.

Table 1: Frequency and proportion of subjects choosing each alternative for problem-S

Alternative	Freq.	Prop. Choosing
①	15	0.146
②	16	0.155
③ (Correct)	54	0.524
④	15	0.146
Blank	3	0.029
Total	103	1.000

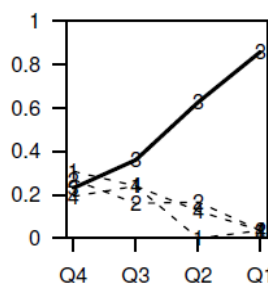


Figure 3: Quantile item response chart for problem-S

As shown in Table 1, the proportion of subjects who answered problem-S correctly is 0.524. Figure 3 shows that the proportion for each group is higher in higher groups with total scores. In the highest scoring group, the proportion who chose the correct answer exceeds 0.8. These results suggest that this problem was quite effective in separating the higher and lower scoring groups in the R1 form, although the subjects had not taught orthographic drawings in the pre-university mathematics curriculum.

To reach the correct answer, a subject has to grasp the holistic shape of the object from the orthographic drawing and read the true length

of sides required to calculate the area of each face. Although the true shape of two faces of the tetrahedron is shown in the front and top views, the true shape of the remaining two faces is not shown in the orthographic drawing. The face with the maximum area is overlapped completely by another face in the drawings and presented as an edge view in the top and right-side views.

According to Figure 3, Alternative 1 was persuasive for the subjects in low scoring groups although only a few of the subjects in high scoring groups chose it. The subject who interprets all of the three triangles shown in the drawing as the true shape could judge alternatives on the basis of comparing the area of the apparent triangles. In this case, Alternative 1 could be chosen as the value of the area of the triangle shown in the front view, or $4 \times 4 \times \frac{1}{2} = 8$. Presumably, most subjects with high mathematics scores were conscious that the triangles shown in the given orthographic drawings are not the true shape of the object face.

Alternatives 2 and 4 were chosen by not only the subjects in low scoring groups but also the subjects in the Q2 group with relatively high scores as shown in Figure 3. Values of both alternatives are larger than that of Alternative 1, which is the value of the face shown as the true shape in the front view. Presumably, the subjects who chose Alternative 2 or 4 were aware of the existence of the face that has a larger area than the face shown as the true shape in the front view and tried to seek the face with the largest area and its value.

The value of Alternative 2 is identical to the area of the face shown in the top and right-side views ($3 \times 4\sqrt{2} \times \frac{1}{2} = 6\sqrt{2}$). It is suggested that the subjects who chose Alternative 2 were not aware of the existence of the face shown as an edge view in the top view. The value of Alternative 4 is not identical to the area of any face of the tetrahedron. It is likely that the subjects who chose Alternative 4 miscalculated the value of the largest area for some reason.

These results suggest the difficulty of problem-S in understanding the shape of the face shown as edge views in orthographic views and whose true shape is not shown in any view.

3.2 Error analysis of problem-L

Table 2 lists the frequency and proportion of subjects choosing each alternative for problem-L. Figure 4 shows quantile item response charts for problem-L. The subjects are divided into four groups (Q1-Q4) with nearly equal numbers in descending order of total score of the R3 form. The proportion of subjects choosing each alternative for each group is plotted except the alternatives chosen by subjects of below 10 % of the total. The proportion of the correct alternative (Alternative 2) is plotted with a thick line.

Table 2: Frequency and proportion of subjects choosing each alternative for problem-L

Alternative	Freq.	Prop. Choosing
①	19	0.260
② (Correct)	52	0.712
③	1	0.014
④	0	0.000
Blank	1	0.014
Total	73	1.000

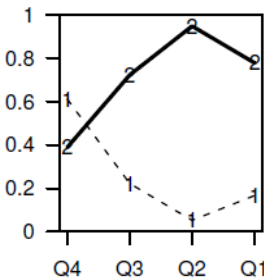


Figure 4: Quantile item response chart for problem-L

According to Table 2, the proportion of subjects who answered problem-L correctly is 0.712. As shown in Figure 4, the proportion choosing correctly is almost 0.4 in the Q4 group and becomes higher in Q3 and Q2 with higher scoring groups. Although the proportion in the Q1 group is below that in the Q2 group, rough fluctuation of the proportion between groups is considered to be caused by a small number of subjects in each group. It was revealed that problem-L also contributed to discriminating between the subjects with high and low scores in the R3 form.

In problem-L, the proportion of subjects who chose the correct answer is almost 0.4 even in the lowest scoring group and exceeds 0.6 in groups other than Q1. A higher proportion chose the correct answer in problem-L than in problem-S, although difficulties of the two problems should not be rashly compared since no common subject was given in both problems.

A subject has to grasp the holistic shape of the object from the orthographic drawing in problem-L. Although the true shape of the base of the quadrangular pyramid is shown in the top view, the true shape of each lateral face is not shown in the given orthographic drawing. To calculate the area of the triangle as each lateral face, the subject has to read the true length of height of the triangle from drawings, whereas the true length of the base of the triangle is already presented in each view.

Wrong answers concentrated on Alternative 1 as shown in Table 2, whereas the proportion choosing each wrong alternative was almost the same as in problem-S. According to Figure 4, the proportion of the subjects who chose Alternative 1 reaches about 0.6 and surpasses that of the subjects who chose the correct answer in the Q4 group. Figure 4 indicates that the proportion choosing Alternative 1 becomes lower in Q3 and Q2 groups, but this is slightly higher in the Q1 group.

The subject who interprets all faces shown in the drawing as the true shape could calculate

each area of the apparent triangles and a square and add them together. In this case, Alternative 1 could be chosen as the value of surface area of the given quadrangular pyramid by using the value $4 \times 2 \times 1/2 = 4$ as the area of each lateral face. It is suggested that such a mistake was often made by subjects with low mathematics ability.

Note that the object presented in the drawings in problem-L has a simple shape with all of the same lateral faces and that the shape of the lateral face can be understood without paying attention to its edge view. In contrast, in problem-S, all faces of the object presented in the drawings are different and understanding the shape of the face shown as edge views plays a key element in finding the correct answer. From these considerations, it is suggested that problem-L is easier than problem-S. Considering that these problems were developed for applicants for liberal arts courses and science courses, respectively, the difference in their difficulties is to be expected.

3.3 Relationship between the performance on the problems in R1 and R3 forms and scores of mathematics in National Center Test

In the investigation, the National Center Test (NCT), which was developed for university entrance examinations in Japan with reference to Japanese curriculum, was also given to the subjects. For mathematics tests, all subjects were given "Mathematics I & A" (hereinafter Math I&A) and "Mathematics II & B" (hereinafter Math II&B) from the NCT regardless of faculties attended. Mathematics II and Mathematics B are subjects designed to be taught to students who have studied Mathematics I and/or Mathematics A. Contents tested in Math II&B are more advanced than those tested in Math I&A. The full score for either test is 100.

Table 3 lists the correlation coefficients between the score of each form made from CEE and the score of the mathematics test from the NCT [1]. The correlation coefficients between

scores in mathematics test forms of both countries were quite high. It is suggested that the mathematics abilities evaluated by the forms of both countries are quite similar, although the constitution of test contents and the test items' response format differ.

Table 3: Correlation coefficients between scores of forms from CEE and scores of mathematics test in NCT (from [1])

		NCT in Japan	
		Math I&A	Math II&B
CEE in Beijing area	R1form (n = 103)	0.71 (**)	0.73 (**)
	R3 form (n = 73)	0.71 (**)	0.74 (**)

(**: $p < 0.01$)

The next interest is how the performance in the problems including orthographic drawings relates to mathematics scores in the NCT. For the subjects given the R1 form including problem-S, the means of Math I&A and Math II&B for subjects who chose the correct answer in problem-S and those for who did not were calculated, and the means of the two groups were t-tested. Table 4 lists the statistics of Math I&A and Math II&B together with the t-values. In both Math I&A and Math II&B, the means of the mathematics test in the NCT significantly differed for the two groups.

In the same way, for the subjects given the R3 form including problem-L, the means of Math I&A and Math II&B were calculated for two groups divided on the basis of the results for problem-L. Table 5 lists the statistics of Math I&A and Math II&B together with the t-values. In both Math I&A and Math II&B, the means of the mathematics test in the NCT significantly differed for the two groups, as was the case for problem-S.

Interestingly, a significant association was found between performance in the problems on three-view orthographic drawings, which are not taught in the mathematics curriculum in

Japan, and the test scores for mathematics developed with reference to the Japanese curriculum.

Table 4: Statistics of mathematics tests in the NCT for correct/wrong in problem-S

(a) Mathematics I & A

		Math I & A		t-value
		Freq.	Mean	
Correct	54	82.0	14.9	$t = 4.649 (**)$
Wrong	49	67.6	16.6	
Total	103	75.1	17.2	

(**: $p < 0.01$)

(b) Mathematics II & B

		Math II & B		t-value
		Freq.	Mean	
Correct	54	85.8	15.0	$t = 3.420 (**)$
Wrong	49	73.6	20.4	
Total	103	80.0	18.7	

(**: $p < 0.01$)

Table 5: Statistics of mathematics tests in the NCT for correct/wrong in problem-L

(a) Mathematics I & A

		Math. I & A		t-value
		Freq.	Mean	
Correct	52	67.1	17.7	$t = 2.396 (*)$
Wrong	21	56.0	18.5	
Total	73	63.9	18.5	

(*: $p < 0.05$)

(b) Mathematics II & B

		Math. II & B		t-value
		Freq.	Mean	
Correct	52	67.6	19.7	$t = 2.696 (**)$
Wrong	21	53.1	23.6	
Total	73	63.5	21.8	

(**: $p < 0.01$)

4. CONCLUSIONS

Multiple-choice mathematics problems, in which an object is presented in a three-view orthographic drawing, were given to Japanese freshmen. A brief explanation of three-view orthographic drawings was added for those who had not studied them in the pre-university mathematics curriculum. It was revealed that the problems contributed to discriminating between examinees with high and low scores in mathematics. The subjects with low mathematics scores tended to choose wrong answers due to misunderstanding a triangle in each orthographic view as a true-size view. In addition, difficulties in understanding the true size of the face presented as an edge view in each orthographic view were suggested. Interestingly, a significant association was found between performance in the problems on three-view orthographic drawings which are not taught in the mathematics curriculum in Japan, and the test scores for mathematics developed with reference to the Japanese curriculum. It is suggested that mathematical abilities developed through Japanese pre-university curriculum play important roles as preparation for understanding of a three-view orthographic drawing.

ACKNOWLEDGMENTS

This work was supported by Japan Society for the Promotion of Science (JSPS) KAKENHI Grant Numbers 24501171 and 23300310. The authors thank all who participated in the investigation.

REFERENCES

- [1] S. Arai, T. Otsu, and H. Miyano. Comparison between mathematics test in university entrance examination in China and mathematics test in National Center Test (*in Japanese*). *Daigaku Nyushi Kenkyu Journal*, 24: 121-126, 2014.
- [2] Y. Zhang, S. Arai, and T. Otsu. National higher education entrance examination in China (*in Japanese*). Research Note

RN-12-02, National Center for University Entrance Examinations, 2012.

ABOUT THE AUTHORS

1. Kumiko Shiina, Ph.D., is a Professor at the National Center for University Entrance Examinations in Japan. Her research interests are evaluation of spatial ability and educational measurement. She can be reached by e-mail at shiina@rd.dnc.ac.jp or postal mail at Komaba 2-19-23, Meguro-ku, Tokyo 153-8501, Japan.
2. Sayaka Arai, Ph.D. is an Assistant Professor at the National Center for University Entrance Examinations in Japan.
3. Tatsuo Otsu, M.A. is a Professor at the National Center for University Entrance Examinations in Japan.

A EULOGY OF AMBIGUITY: BETWEEN VISUAL PERCEPTION AND QUANTUM MECHANICS

Giuseppe CAGLIOTI and Luigi COCCHIARELLA

Politecnico di Milano, Italy

Ambiguity is hardly a failure of rationality, it is a place where art and science can meet and talk

Roald Hoffmann, Private communication, 1997

ABSTRACT: Assuming ambiguity as a coexistence and/or coalescence of two incompatible or contradictory aspects in the same reality, may a eulogy of ambiguity make any sense in general? Moreover, what kind of connection can relate the human world of visual perception and the physical world of quantum mechanics?

Concerning the first question, from a scientific point of view, we know that in nature ambiguity occurs at some critical points in the physical transformations, for instance the water at the freezing temperature is liquid and at the same time solid, while from an anthropological point of view we could mention the same nature of deities, often ambiguously considered half human and half divine, and we could also mention the powerful role of ambiguity in the creative stages of an activity, be it science, art or technique, providing room for new and sometimes unexpected hypotheses, or the benefits of the contaminations among human traditions, disciplinary fields, and so forth. Anyway, in spite of these and of many other evidences, in our western culture the role of the ambiguity has been gradually underestimated, especially by the scientists, at least since the Aristotle's principle of *tertium non datur*.

Concerning the second question, since the early times of quantum mechanics Werner Heisenberg pointed out the inadequacy of the ordinary language to express the *Weltanschauung* of the microscopic world of the atoms, and surprisingly it seems that Geometry and Graphics could offer convincing strategies to realize ambiguous visual images at the same time representing ambiguous physical states. More precisely, this work compares the *bi-stable* visual effect occurring during the perception of some pictorial configurations (matter of Kinetic Art), to the evolution of the *bi-stable* status of a molecular quantum structure perturbed during a spectroscopic measurement (matter of Quantum Mechanics). In both cases a sequence of quasi-periodic alternations can take place, so that *space* and *time* enter into question together with *energy*, and this latter can be measured, namely in the physical structure and probably in the brain activity. Dealing with these transformations, the notions of *order* (as *correlation*) and *symmetry* (as *invariance*) and their interplay will be discussed from a wider point of view and distinguished according to their nature and effects, while on the other side the links between the *emotional* and the *cognitive* power of images should appear more tangible in this new light, and consistent with the results of the neurosciences.

Some examples will be shown, reconfirming the 'double nature' of the *image*, namely its *figurative* (graphic) and *metrical* (geometric) identity, maybe the deeper reason of its versatile role across science and art, and of its intrinsic ambiguity.

Keywords: ambiguity, symmetry reduction, order, visual perception, quantum mechanics, spectroscopic measurements.

1. INTRODUCTION

The ambiguity, despite its negative ethical concept, is an existential human value [3-9].

Ambiguity can be defined as the coexistence and/or coalescence of two incompatible or contradictory aspects in the same reality. In Nature the ambiguity occurs systematically:

a) at the critical point of the structural transformations of physical systems near the thermodynamic equilibrium. Think, for instance, to the water at the freezing temperature is a liquid and at the same time a solid;

b) at the critical point of the instabilities of non-linear dynamic systems removed away from equilibrium by incoming fluxes of resources – materials, energy, information. Think, for instance, to a liquid heated from below, precisely at the onset of the convective motion triggered by the upward Archimedean buoyancy force (see Fig. 1).



Figure 1: The Giant Causeway in Ireland and its laboratory analogue (Courtesy of Manuel G. Velarde).

Since the Aristotle's principle of *tertium non datur* [10], ambiguity has been underestimated in our western culture [11]. However, the ambiguity, also in our mind, occurs continuously, eg. at the onset of visual thinking during the process of perception and at the critical state of any decisional process.

After some considerations on the role of the symmetry reduction in the process of perception we highlight the cognitive power of visual perception of the ambiguous images of the kinetic art. Somewhat surprisingly the process of perception of these images can be described in terms of a quantum-like evolution of our mental state [7-9, 14, 22]. We aim to show that quantum mechanics can help to comprehend the process of perception of ambiguous images, and vice versa that the process of perception of these images can help to visualize basic concepts of quantum mechanics (*perception* and *comprehension* cooperate for the acquisition and dissemination of knowledge: remind that in Latin languages the etymological root *capio* and the Italian *prendo* mean: *I take*).

2. PERCEPTION AS THE OUTCOME OF A MENTAL DYNAMIC INSTABILITY

The visual perception of an image can be envisaged as an irreversible ordering process developing in our mind.

During perception the mental state of the internalized image evolves. Its evolution is activated by the fraction of the mental energy employed for the pattern recognition and is controlled by curiosity and attention [19-20]. The internalized image evolves in our mind through a sort of non-equilibrium dynamic instability, from an initially balanced, disordered, statistically "symmetric" state produced by a heap of undifferentiated responses to the stimuli, to an unbalanced, ordered state where the stimuli lead to a coherent, holistic entity; from an initially meaningless structure to the thought or the visual thinking [1]. At the critical point of this dynamic instability, the *indiscernability* (a characteristic of the symmetry) and the *correlability* (a characteristic of the order and the organization) coalesce and coexist ambiguously.

At this critical point, a fluctuation and it is catastrophe: a reduction of the “statistical” symmetry occurs and disorder is removed, while the visual thinking grows and the internalized image merges into coincidence with an archetype genetically or culturally formed and stored in our memory. We can experience vividly this complex process during the perception of figure 2.

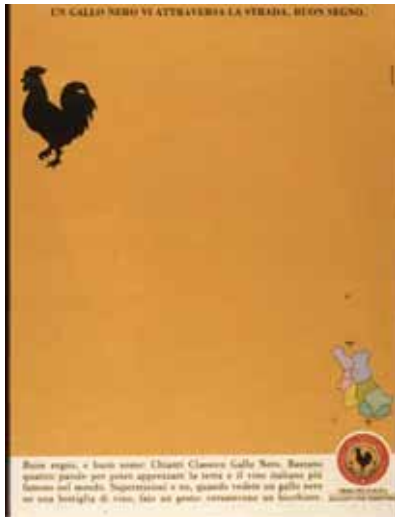


Figure 2: Advertising page of a newspaper.

The advertising page shown in the figure 2 is almost entirely occupied by a homogeneous, meaningless background, characterized by the exchange symmetry: if two graphic components of the background are interchanged, no change can be discerned. Our eyes are eager to extract information, they run instinctively toward the contour of the logo – the rooster – where the (exchange) symmetry is broken.

Also in Fig. 3 our attention is attracted by the contours of the black spots, which exhibit a breaking of the exchange symmetry. Since these spots are numerous, the process of perception of this image requires more time and more mental energy than figure 2.



Figure 3: An image by R.C. James (1966).

An intense, unconscious mental work is required in order to scan mentally the contours of the numerous black spots where the exchange symmetry breaks. When the disorder of a yet meaningless internalized figure and the order of an already organized visual thinking merge together, a critical, ambiguous mental state is reached. A small fluctuation and it is catastrophe: a sudden global symmetry reduction occurs and the internalized figure falls into coincidence with the archetype of the dog, impressed in our brain.

3. VISUAL THINKING OF AMBIGUOUS IMAGES AND ITS QUANTUM MECHANICS ANALOGUE

We focus now on paintings generating optical “illusions”, such as those by M. C. Escher, or the ambiguous structures of the kinetic art by V. Vasarely and F. Grignani.

Also for these paintings the end point of the perceptive process is achieved by the onset of visual thinking. But here the archetype into which the interiorized image merges as visual thinking is formed doesn’t correspond to the idea of a static archetype. The situation is more intriguing because visual thinking doesn’t reach a unique, stable configuration but oscillates between two (and sometimes more) alternating patterns. Psychologists call this phenomenon *multistability in perception* [7]. However the fact that at the end of the perceptive process no

stability is reached doesn't imply that these abstract multistable structures of the kinetic art are less meaningful and less adherent to reality than those depicting static objects. The lack of perceptive stability doesn't imply that these structures are "impossible" or misleading. On the contrary we purport here that it would be misleading to classify the perceptive permutations of ambiguous figures as an optical *illusion* [9]. In this perspective, though perhaps paradoxically, the title given by M. C. Escher to one of his most famous paintings is *Concave 'and' convex* and not *Concave 'or' convex* [5]. Correspondingly, in Nature, at the critical point of the diffusion \rightarrow convection dynamic instability the liquid is still *and*, *at the same time* is in motion [24, 27]. Furthermore at the critical point of the paramagnetic \rightarrow ferromagnetic equilibrium phase transition of magnetite, the magnetite is paramagnetic *and at the same time* is ferromagnetic.

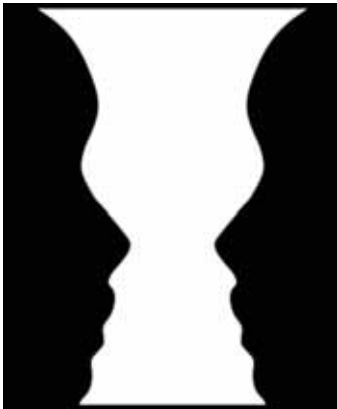


Figure 4: A vase 'and' two profiles (1921), by Edgar Rubin

The Danish psychologist Edgar Rubin, the author of the figure 4, was a close friend of college of Niels Bohr, one of the founders of the quantum mechanics, and most likely influenced the interpretation proposed by Bohr of the dual behaviour – particle-like and wave-like – of elementary object such as electrons, pho-

tons, ... According to Bohr, "...evidence obtained under different experimental conditions cannot be comprehended within a single picture, but must be regarded as complementary in the sense that only the totality of the phenomena exhausts the possible information about the objects." [2]. Let us now consider one of the simplest structures of the class of gestaltic, bistable, modular, resonating ambiguous figures (Fig.4). The archetype corresponding to this image evokes the very basic mental mechanism underlying the dynamics of ambiguity, The dynamics of ambiguity, perceived while realizing that the modular structure of Fig. 5 is actually resonating, is formally analogous to that occurring during the spectroscopic measurement of a molecular quantum structure [3].

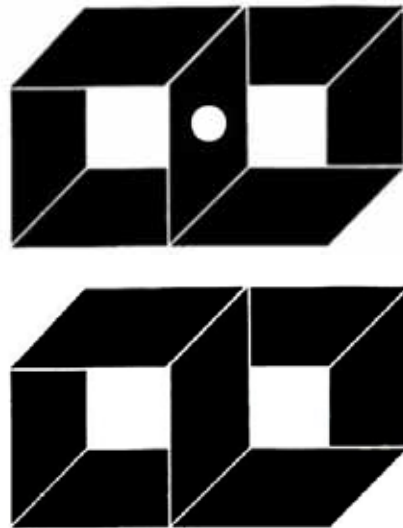


Figure 5: An example of bistable figure.

At first this figure could be envisaged superficially as a two-dimensional structure exhibiting a center of symmetry. Keep looking at it as passively as possible, so as *not to perturb* the figure, which by construction is confined in the two-dimensional plane. Its central region

around the center of symmetry could be described in two ways:

a) as belonging 50% to the modulus at left and 50% to the modulus at right in a center-symmetric stationary state;

and, paradoxically, at the same time

b) as not belonging neither to the modulus at left nor to the modulus at right in a center-antisymmetric stationary state [3].

These two descriptions, though quite acceptable if considered separately, are incompatible if attributed simultaneously – as they should in this case – to the same reality. Indeed we react instinctively to the “absurdity” of the situation and hasten to remove the ambiguity built in the figure: suddenly we spontaneously let the two-dimensional figure invade the three-dimensional space and assign the central region to the right hand cubic modulus or to the left hand cubic modulus. However, at this critical point, our choice implies that the center of symmetry disappears (Fig 5, above): a symmetry reduction occurs (Fig. 5, below). Thereafter, visual thinking cannot remain fixed in any of these two positions: soon an organized sequence of quasi periodic perceptive alternations of right → left → right ... prospects sets in.

Curiously, yet with reference to Fig. 5, the process of *perception* leading to the visual thinking turns out to share this ambiguous feature with a typical process of *measurement* in quantum mechanics. We refer e.g. to the so called *charge transfer spectrum* of a homonuclear diatomic molecule, well known to the molecular spectroscopists of the Fifties [21]. A homonuclear diatomic molecule (e.g. the hydrogen molecule, H_2) is *described* by the positions of its valence electrons.

The possible distributions of the positions of these electrons in the space are classified in terms of stationary, essentially time independent molecular states.

The most important ones are:

- a) the fundamental gerade center-symmetric bonding state (Fig. 5, at bottom, right), and,
- b) the ungerade center-antisymmetric antibonding state (Fig. 5, at bottom, left).

An electronic energy level is associated to each of these stationary states. However *to describe* is not equivalent to *to measure*: the description ignores the time dimension, while the measure implies an irreversible action hopefully leading to an information; and by definition any action requires time as well as energy (In quantum physics the action is a fundamental concept; it is expressed by the product of an energy by a time and is measured in units of the Planck's constant, $h = 6,6 \cdot 10^{-34} \text{ J}\cdot\text{s}$). In other words in order to pass from an aloof description to an active measurement one has to enter into the time dimension.

In practice the measurement of the difference between the energy levels associated to the fundamental gerade center-symmetric bonding state and the ungerade center-antisymmetric antibonding state, $2K$, is accomplished setting the molecule in the flux of an electromagnetic radiation. If the frequency ν of the radiation corresponds to the above energy difference (so that the condition for the spectroscopic resonance $h\nu = 2K$ is satisfied), a resonant absorption occurs: the electric component of the radiation field hooks the valence electron and drives it from the atom at the right hand to the atom at the left hand and vice versa, so producing a resonant exchange of the electron itself between the attracting positive nuclei. This resonant exchange closely reminds the beats of the resonating strings of the *viola d'amore*, or, in physical space, the formation of the Moiré pattern generated by the superposition of two gratings of similar periods.

The spectroscopist ends up finding the electron where the resonant electromagnetic field drives it while searching for it. The paradoxical (?) features of this situation can be experienced visually focusing on figure 6 (center, bottom) or, equivalently on its detail (Fig. 5).

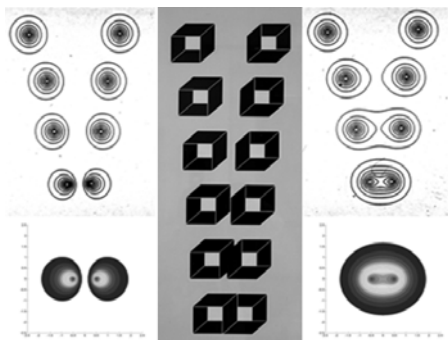


Figure 6: The quantum mechanical representation of the hydrogen molecule ion.

In the figure 6, the antibonding (at left) and bonding (at right) stationary states of the electron are shown for decreasing distances between the attracting positive charges of the protons. The probability density of finding the electron in the unit length along the molecular axis is plotted for each one of the above two states. The *Condensazione Grafica* by Franco Grignani (in the middle) proposes a dynamic representation of the molecule. The white squares of the 'condensed' ambiguous figure in the bottom represent the positively charged protons attracting the negative *black* electronic charge. The dynamic perceptive inversions, produced by the visual thinking, remind the resonating oscillations of the electronic charge of the hydrogen molecule H_2 during the measurement of the *charge transfer spectroscopy*. When the frequency of the electromagnetic radiation impinging on the molecule resonates on the gap between the energy levels of the H_2 bonding and antibonding molecular states, the electric component of the electromagnetic field hooks the electron and transfers it periodically from one Hydrogen atom to the other one.

In this perspective the process of spectroscopic measurement and the perceptive process leading to the optical illusion (but are we still thinking that it is really an *illusion*?) share the same paradoxical (?) features: although both processes are time dependent and *irreversible* by nature, they can be understood only in terms

of the *stationarity* of the time-independent center-symmetric and center-antisymmetric states. Every teenager, unaware of the intricacies of quantum mechanics but capable to accomplish the dynamic perception of an ambiguous figure such as those presented above, can grasp the paradox of the incompatibility between the stationarity of the states in whose terms a quantum structure must be described and the irreversibility of the process of measurement [26].

During the early days of quantum mechanics Werner Heisenberg felt the inadequacy of everyday language to express the emerging *Weltanschauung* of the microscopic world of the atoms. His remarks (On Quantum Physics, 1930) are transcribed below:

"It is not surprising that our language should be incapable of describing the processes occurring within the atoms, for, as has been remarked, it was invented to describe the experiences of daily life, and these consist only of processes involving exceedingly large numbers of atoms. Furthermore, it is very difficult to modify our language so that it will be able to describe these atomic processes, for words can only describe things of which we can form mental pictures, and this ability, too, is a result of daily experience".

The analogy between the measurement of quantum structures and the process of *perception* induces us to think that the difficulty denounced by Heisenberg can be overcome if one accepts the fact that perception can be utilized as an useful tool for the achievement and dissemination of knowledge [26].

4. BACK TO THE VISION BACKSTAGE

The fact is that the ambiguity affects the basic physiologic processes of the visual perception. As the vision is based on a combination of (optical) *projection* and (intellectual) *cognition*, it is interesting to see how both these phenomena show ambiguous aspects.

From a *projective point of view*, we know that the space of our everyday life can approximately be considered Cartesian, while its optical projection on the retina follows a 'polarized' conic model depending on the sight point.

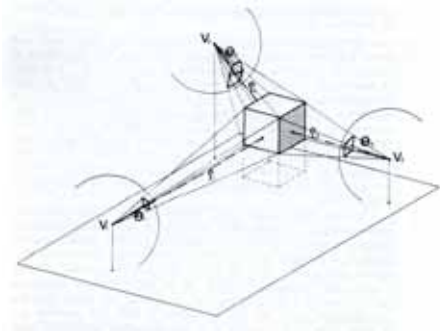


Figure 7: true space and its conic projections on a curved surface as a projective models of the visual perception [16].

The figure 7 refers to a theoretical model, but the shape of the images on the retina has been actually detected through experiments carried on since the XVII Century. The first proofs concerning how the light casts on the retina have been found by Christoph Scheiner (1619) and René Descartes (1637), so confirming that the light runs *towards* the eye, as Johannes Kepler noticed (1604), and not vice versa, as people wrongly believed since the Euclide's time. More recently, new experiments have been proposed by using new technologies to keep humid the ocular bulbs, like the one discussed by Maurice Henry Pirenne [25]. In this case, an ocular bulb taken from an albino rabbit has been used because of the transparency of its choroid. Then, in a dark room, two parallel series of aligned light-spots have been located in front of the mentioned organ during the experiment. Entering the pupil and passing through the vitreous body, light rays reached the back of the sclera, forming there the projected images of the spots. The whole configuration of the resulting dotted figure follows the laws of the curvilinear perspective, as the pictures of the experiment clearly show (Fig. 8). The set has been arranged in order to test the projective cast in different situations, moving the two series of light-spots around the ocular bulb (Fig. 9).

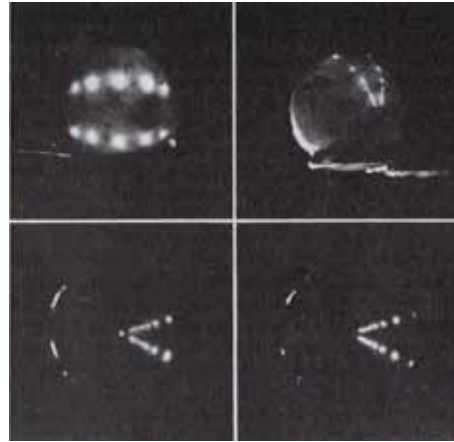


Figure 8: pictures of the image of aligned light-spots appearing on the surface of the retina of an albino rabbit [25].

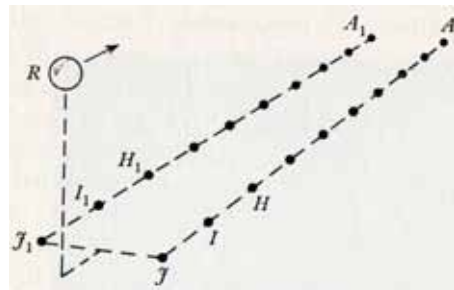


Figure 9: A set arrangement where light-spots are parallel to the visual axis [25].

The experiment clearly shows how the foreshortenings work on the projective surface of the retina. Lines are projected as points or curves and parallel lines as curves meeting at one or two points, according to the setup. The linear structure of the true three-dimensional space is turned into a non-linear bi-dimensional visual field, whose curvature depends on the geometrical feature of the ocular bulb. We know, in fact, that the real configuration of the light signals displayed on the retina is strongly distorted (Fig. 10).

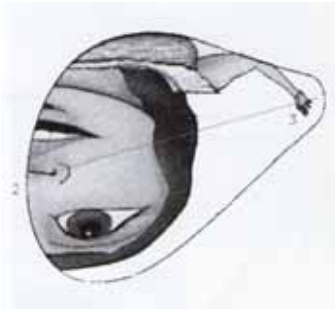


Figure 10: A projection on the cortical surface of the ocular bulb [15].

Agostino De Rosa pointed out that the cortical projections onto the ocular surface are actually *anamorphic* images [15]. Moreover, our visual perception is based on a binocular system, where series of couples of distorted images are combined thanks to our brain during the visual process. Since we see with two eyes, left and right cortical images are partially different one from the other, according to the distance between the two sight points (Fig. 11). Therefore, the first ambiguity concerns the continuous switch from 2D images perception to 3D mental reconstruction.



Figure 11: Combination of the cortical projections in the visual perception [15].

Sometimes reconstructions are also affected by the projective ambiguity of the images.

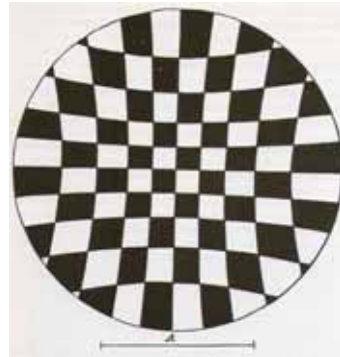


Figure 12: Test image to see the variations of the apparent curvatures when changing the viewing distance. H.von Helmholtz (1866)[16].

In fact, some optical ambiguities have to be taken into account, like the instable images of the curves, sometimes appearing as straight lines due to their position or to the curvature of the retina. For example, the diverging curves forming the tiles in the well-known image by Hermann von Helmholtz (1866), shown in the figure 12, tend to appear like straight lines when we look at the center of the image from a very short distance, and several other examples could be mentioned. Last, sometimes no images occur at all, since our eyes have a *blind point* where the optic nerve meets the retina. Therefore our brain has to provide a mental completion of this 'hole' in order to avoid perceptual discontinuities. The effect of this blind point can be anyway detected with a simple experiment, using the two disks appearing in the figure below (Fig.13).



Figure 13: Test image to detect the existence of a blind point onto the retina [16].

Looking at the figure 13 from a very short distance (less than 25 cm in this case), close the left eye and look at the left disk, then adjust your sight until the disk disappears. This happens when the light rays projecting the image cast on the blind area of the retina. The experiment can be repeated keeping closed the right eye while looking at the right disk. This is a case of extreme ambiguity, when the alternative is not between ‘something’ and ‘something else’, but between ‘something’ and ‘nothing’. It is clear at this point that we do not see the images on the retina, since they are signals that need to be processed by the brain.

From a *cognitive point of view*, our second issue, we actually do not see our retina neither our mental processes, but the *result* of their interaction. Moreover, since our visual perception requires a great deal of energy, brain uses some strategies to optimize the number of visual information to process. Rudolph Arnheim has well explained this aspect, showing that we usually take advantage of some familiar perceptual patterns in our everyday life. Tomàs Maldonado has hypothesized that images can be simplified until to their *iconic unit*, or better, until to reach the minimal overall structure that make them still recognizable at a glance. To explain this idea, he proposed the neologism *icon-eme* (in Italian *iconema*), meaning something very similar to a phoneme or a morpheme in the field of linguistics [13][23]. We think that this theory is valid not only in relation to the objects, but also in relation to our visual patterns, that are basically optical-perspective.

In fact, looking at the figure 14, a structural contradiction is immediately detected, that pushes us to gradually overestimate the real graphic size of the black figures of the human beings on the right side of the illustration, in spite of the fact that they have the same dimension. The apparent dimensional inversion testifies that we do not see the true shapes, being us strongly influenced by the more familiar *icon-eme* of our optical patterns, that are largely prevailing because we usually move in a true space that is perceived according to their rules.



Figure 14: Perspective size changes [16].

Besides the contradiction between two or more patterns, sometimes even one single pattern can lead into ambiguous perception, like in the cases represented in the figure 15, where the orientation of the deep lines makes difficult to understand which parts of the object are close and which ones are far, as noticed by Roberto De Rubertis [16]. The figures are visually bistable, which means, appearing in two alternative positions. Although we expect this effect in the parallel projections (A), it is interesting to note that the instability is not mitigated in the perspective representations (B).

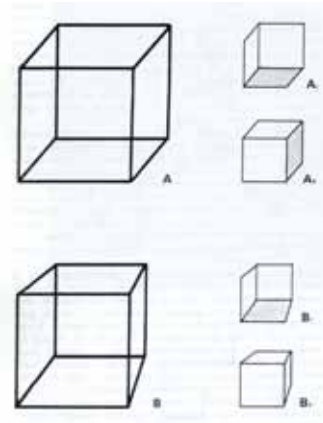


Figure 15: Bistable visual patterns [16].

This situation is the most directly related to the quantum mechanics phenomena described in the first part of the paper. From a strictly projective point of view, which means if we do not consider the lenses, we know that a single projected image as well as a single eye vision does not preserve the one-to-one correspondence with the original objects. Then, we can not exactly recognize the real shape of a space basing on one picture only, as well as we can hardly distinguish a painted or a three-dimensional *thrompe l'œil* when looking at it from the designated illusory sight point. In the figures 16 we see two *stenopeic* pictures, taken with a *stenoscope*, that is a special photo-camera without lenses, in order to avoid any curvature. They show the perspective Gallery of Palazzo Spada in Rome, by Francesco Borromini. Observing the picture on the left, taken from the right point of view, no one could say that the real gallery is not 'cylindrical', but 'conical' and much more short than it appears, perfectly realized as a three-dimensional perspective projection. It is interesting to compare this image with the picture on the right, taken from a wrong point, namely from the end of the gallery, producing a very different visual effect.



Figure 16: Perspective Gallery of Palazzo Spada in Rome (1652). Views from the designated point (main entrance, left) and from another point (opposite entrance, right)[25].

More generally, in spite of the more refined spatial estimations we make when observing an illusory image, object or space 'from life', with

both the eyes, then taking advantage of the *collineation* of the visual rays, the illusion still tends to successfully work, thanks to the combined effects of the feature of the object and of our optical-perspective perceptual structures.

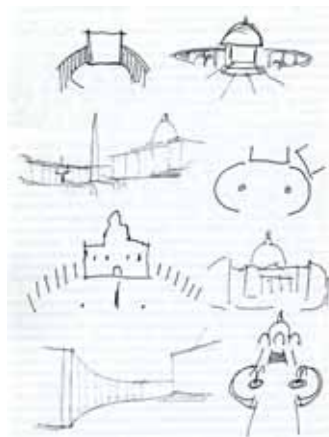


Figure 18: St. Peter Church in Rome, strange drawn images from memory [16].

Other interesting instabilities happen when dealing with memorized images, like in the drawings included in the figure 18, realized by people asked to describe the Saint Peter's church in Rome; particularly intriguing is the drawing on the lower right side, where the front of the church looks like much more similar to the Saint Mark's basilica in Venice than to the Saint Peter's church in Rome [16].

In conclusion, we agree with the neurobiologist Semir Zeki's opinion, remarking the relationships between *ambiguity* and *consciousness*, and how the first one would not be possible without the latter. He thinks that *ambiguity* is a fundamental part of how the brain works, at least the visual part of it. In other words, it offers rich opportunities to the knowledge development, providing room for alternative and/or coexisting cognitive hypotheses during the perceptual as well as the creative processes, and human beings normally experience and use it either in science or in art.

5. CONCLUSIONS

The introspective analysis of the process of visual perception of intriguing ambiguous paintings of the optical art can be exploited as a tool enabling to disclose complex concepts of the quantum physics and the structure of matter, bypassing the intellectual efforts usually needed in order to understand them through the traditional learning activity. These concepts contribute to shape our *Weltanschauung*. And yet we have shown that they are intimately connected with ambiguity, optical illusions and paradoxes: with features usually attributed to the world of art rather than to science. But art is a form of knowledge, and the artist's intuition often precedes scientific formalizations by centuries or even millennia [1].

REFERENCES

- [1] R. Arnheim. Visual Thinking. Regents of the University of California, Berkeley, 1969.
- [2] N. Bohr. Discussion with Einstein on Epistemological Problems in Atomic Physics. In P.A. Schilpp, P. A. (et al.). Albert Einstein: Philosopher-Scientist. The Library of Living Philosophers Inc., Evanston, Illinois, p. 210, end of first passage, 1949.
- [3] G. Caglioti. The Dynamics of Ambiguity. Springer, Berlin, Heidelberg, New York, 1992 [available also online]; *Symmetriebrechung und Wahrnehmung*, with a preface by H. Haken. Vieweg, Braunschweig, 1990; Russian Edition, with a preface by I. Prigogine. MIR, Moscow, 1997; Japanese Edition, with a preface by K. Husimi. Kodansha, Tokyo, 1997; Italian original: *Simmetrie infrante nella scienza e nell'arte*, con Prefazione di P. Fenoglio. CLUP, Milano, 1983; 2nd Edition, with a preface by H. Haken. CittàStudi, Milano, 1994.
- [4] G. Caglioti. Eidos e Psiche. Struttura della materia e dinamica dell'immagine. Ilisso, Nuoro, 1995; Japanese Edition, Yakuyosha Co., 2001.
- [5] G. Caglioti. The World of Escher and Physics. In *M.C. Escher: Art and Science*, (H.S.M.Coxeter, M.Emmer, R.Penrose and M.L.Teuber eds), North Holland, Amsterdam, 287 1986.
- [6] G. Caglioti. Ambiguity in Art and Science. In *Proceedings of the Workshop on Art and Science. (ENEA/WAAS, Vinci, 11-13 December, 1992)*, pages 63-74. World Futures 40, 1994.
- [7] G. Caglioti. Perception of Ambiguous Figures: A Qualitative Model Based on Synergetics and Quantum Mechanics, (P. Kruse and M. Stadler, eds). In *Ambiguity in Mind and Nature: Multistable Cognitive Phenomena*, p. 463. Springer Verlag, Berlin, Heidelberg, New York, 1995.
- [8] G. Caglioti. From perception to thought, A nonequilibrium dynamic instability implying symmetry breaking. In *Katachi U Symmetry*, Springer-Verlag, Tokyo, 1995.
- [9] G. Caglioti. Optical art: "illusions" and paradoxes of symmetry and quantum mechanics. In *Symmetry 2000* – part 2, I. Hargittai and T. C. Laurent, Eds, pp. 457-466. Portland Press, London, 2002.
- [10] G. Caglioti. The tertium non datur, in Aristotle's Logic and in Physics. *J. Mech. Behaviour of Materials*, 5, 3: 217-223, 1994.
- [11] G. Caglioti. "Ambiguity". An Entry in the *New dictionary of the History of Ideas*, Scot Peacock, Charles Scribner's Sons Editor, pp. 53-55, Farmington Hills, MI 48331-3535, 2004.
- [12] G. Caglioti. Cognitive Power of Visual Images. In L. Cocchiarella (editor), *The Visual Language of Technique: Heritage and Expectations in Research*. Vol. 2, Springer (2014, to appear).
- [13] L. Cocchiarella. *Fra Disegno e Design: temi forme codici esperienze*. Città Studi, Milano, 2009.
- [14] E. Conte, A.Y. Khrennikov, O. Todarello, A. Federici, L. Mendolicchio and J.P. Zbi-

- lut. Mental states follow quantum mechanics during perception and cognition of ambiguous figures. *Open Systems and Information Dynamics*, 16, 1, 85-100, 2009.
- [15] A. De Rosa, G. D'Acunto. La vertigine dello sguardo. Tre saggi sulla rappresentazione anamorfica. Cafoscarina, Venezia, 2002.
- [16] R. De Rubertis. Progetto e Percezione. Officina Edizioni, Roma, 1971.
- [17] G. Fano, Correzioni ed illusioni ottiche in architettura. Dedalo libri, Bari, 1979.
- [18] R.L. Gregory, J.G. Wallace. Recovery from Early Blindness. *Experimental Psychology Society Monograph*, 2: 1963.
- [19] H. Haken, edr. Pattern Formation by Dynamic Systems and Pattern Recognition. In *Proceedings of the International Symposium on Synergetics at Schloss Elmau, Bavaria, April 30-May 5, 1979*. Springer, Berlin, Heidelberg, New York, 1979.
- [20] H. Haken. Synergetics, An Introduction. 2nd ed., Springer, Berlin, Heidelberg, New York, 1983.
- [21] G. Herzberg. Molecular Spectra and Molecular Structure I - Spectra of Diatomic Molecules. Van Nostrand Reinhold, New York, 1950.
- [22] A. Khrennikov. Ubiquitous Quantum Structure: From Psychology To Finance. Spriger, 2010.
- [23] T. Maldonado. Reale e Virtuale. Feltrinelli, Milano, 1992.
- [24] G. Nicolis, I. Prigogine. Self-organization in Non-equilibrium Systems. Wiley, New York, 1977.
- [25] M.H. Pirenne. Optics, Painting & Photography. Cambridge University Press, Cambridge, 1970.
- [26] I. Prigogine, I. Stengers. La nouvelle Alliance. Scientia 112, pp. 287-332 and 617-653 (Available only in French), 1977.
- [27] M.G. Velarde, C. Normand. Convection. In *Sci. Amer.* 243 (1), 1978.
- [28] Zeki S. Neurologia dell'ambiguità. In: Lucignani G., Pinotti A., Raffaello Cortina, Milano, 2007, pp. 83-119. Original version: Zeki S. The neurology of ambiguity. *Consciousness and Cognition*, 13-1, 2004, pp.173.196.

ABOUT THE AUTHORS

1. Giuseppe Caglioti: Physics degree (Roma, 1953), Diplomas in *Nuclear Physics* (Roma), and of *School of Nuclear Science and Engineering* (ANL, Illinois, USA). Research on nuclear physics, neutron spectroscopy, structure and mechanical performance of materials: in Roma, in ANL, USA, in AECL, Chalk River, Canada and with ENEA in the JRC of the EU, Ispra, Italy. Intermittently, in BNL, USA and in the University of Stony Brook, NY, USA. Research is now focused on the PiezoMusicColor Project, a natural form of technologic art aiming to a synaesthetic enjoyment of music and sounds also by blind, deaf and hard of hearing persons (see www.piezomusicolor.it). Author/co-author of several books, two patents, about 130 scientific publications, and 50 papers on correlations between human and natural sciences. Since 1970 he has been full professor of solid state physics and docent of Aesthetics – Scientific Components of Harmony and Beauty at the Politecnico di Milano. At present, Emeritus Professor, Politecnico di Milano, he can be reached by email at: giuseppe.caglioti@polimi.it.
2. Luigi Cocchiarella: Architect, Phd, since 1999 he is Researcher (Assistant Professor) at the Politecnico di Milano, where he has been teaching at the Schools of Architettura e Società, and of Design. His research interests are in the field of Geometry and Graphics, their techniques and history, including computer graphics and graphic education. Member of the Department Board of the Department of Architecture and Urban Studies (DASTU) he can be reached by email at: luigi.cocchiarella@polimi.it.

EVALUATION OF PASSING PERFORMANCE OF SELF-DRIVEN PARTICLE THROUGH BUILDING

Takuya MATSUMOTO, Kensuke YASUFUKU and Hirokazu ABE
Osaka University, Japan

ABSTRACT: In building or urban planning, we must consider the passing performance of various objects such as people or cars. To design comfortable, safe, and secure space, it is important to evaluate the permeability of the space. This research proposes an evaluation index and develops a simulation method of the passing performance of autonomous objects by using self-driven particles and lines of sight. We define the resistance value as an evaluation index of the space permeability and measure the resistance values when self-driven particles or lines of sight pass through varying arrangement patterns. We classify the measurements into four groups and clarify the characteristics of each group.

Keywords: Design argument, passing performance, permeability, self-driven particle, processing.

1. INTRODUCTION

In building or urban planning, we must consider the passing performance of various objects. As an example, much research in architectural environmental engineering has studied the passing performance of a natural environment, which includes light, airflow, and sound [1]. However, evaluation indices of the passing performance of objects that move while interacting (such as people and cars) have not yet been established. Alice Coleman, a well-known geographer, used the word “permeability” to evaluate the route selectivity of a pedestrian or the visibility on an urban street [2]. To design comfortable, safe, and secure space, it is important that we evaluate the permeability of space.

In recent years, computer simulations of human or car movements have advanced. By using simulations of self-driven particles or a multi-agent system, we can model the detailed movements of humans and cars. Moreover, the application range of simulations is becoming diverse. However, the simulation method has been mainly used to evaluate evacuation safety

or to analyze the cause of traffic congestion, and it is not much utilized to evaluate the passing performance of space.

The purpose of this study is to propose an evaluation index of passing performance and to develop a method of evaluation. This research proposes an evaluation index and develops a simulation method of the passing performance of autonomous objects by using self-driven particles.

2. METHOD

2.1 Developing a system that evaluates passing performance

We develop a system that evaluates the passing performance of a space when objects, particles, or waves pass the space. Passing objects, particles, or waves can be people, cars, winds, sounds, lights, lines of sight, or radio waves. The goal of this research is to evaluate the permeability of architectural or urban space. To evaluate permeability, some indices are necessary: indices of the route selectivity of pedestrians, visibility in urban streets, continuity between urban streets and buildings, and visibility from buildings to an urban street [3]. In this

study, we focus on indices of the route selectivity of pedestrians and visibility in urban streets. As a first step of measurement for using these indices, we model two types of passing object: (1) a self-driven particle as a simplification of the movement of a pedestrian who selects a path and (2) the line of sight as a simplification of the sights of many people on an urban street.

2.2 Self-driven particles

Because living things move according to their own free will, they do not follow the law of inertia. Particles like these are called “self-driven particles” to distinguish them from particles that follow Newtonian mechanics [4]. For the self-driven particles in this study, we use the Boids algorithm, which has been used to express bird or fish flocking. Boids applies the following three rules [5].

- Separation: steering to avoid crowding local flockmates
- Alignment: steering towards the average heading of local flockmates
- Cohesion: steering to move toward the average position (center of mass) of local flockmates.

These rules are shown in Figure 1.

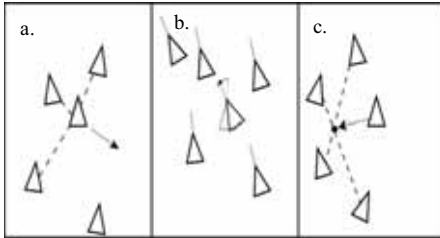


Figure 1: Three rules of Boids

In this study, we added three more rules to model flock movement.

- Particles go in a certain direction. In a certain direction, particles move at 4000 mm/s^2 .
- If a particle goes near an obstacle, the particle

moves away from the obstacle at $1/d \times 2 \text{ mm/s}^2$. (Distance from obstacle is d .)

- If a particle hits an obstacle, the particle rebounds.

2.3 Lines of sight

As a second type of passing object, we model the “line of sight”. The purpose of “line of sight” is to evaluate the passing performance of a line of sight in architecture or in a city. Thus, we evaluate the visibility when we look in one direction. For example, the line of sight shows how much of a city block a person can see from a main street.

2.4 Evaluation method of passing performance

To quantify passing performance when self-driven particles or lines of sight pass through a space, we define the “resistance value”. Resistance value of self-driven particles is Rs , line of sight is Rer and reflected line of sight is Res .

The resistance value of a self-driven particle is calculated by using the time required for self-driven particles to pass through an area of space. Under the following two conditions, we measure the time T required for passing an area of space.

- There are obstacles. (Time is $T_1(i)$.)
- There are no obstacles. (Time is $T_0(i)$.)

Here, i is the i th self-driven particle of N self-driven particles. We calculate the average passing speed of each self-driven particle, where L is the distance from one end to the other end of the area of space.

$$V_0(i) = \frac{L}{T_0(i)}$$

$$V_1(i) = \frac{L}{T_1(i)}$$

We define that the speed of a self-driven particle that stops on the way is 0. We calculate the average speed from $V_0(0)$ to $V_0(N)$ and from $V_1(0)$ to $V_1(N)$. Formula (1) shows the resistance value Rs .

$$R_s = \frac{\sum_{i=1}^N V_0(i)}{\sum_{i=1}^N V_1(i)} \quad (1)$$

This equation shows that the nearer a resistance value is, the higher the passing performance is, and that the larger the resistance value is, the lower the passing performance is.

The resistance value of the line of sight shows the passing performance of an area of space when N lines of sight go straight in one direction. Some lines of sight are blocked by obstacles, but we only measure the number of lines of sight (N_l) that reach the end of the area. Formula (2) shows the resistance value R_e .

$$R_e = \frac{N}{N_l} \quad (2)$$

Furthermore, we can add the line of sight to a reflection. If the line of sight intersects the obstacle, the line of sight is reflected. In this case, the strength of the line of sight is reduced by half every time the line of sight is reflected. We calculate the resistance value R_{er} from strength $F(i)$ of the line of sight that reaches the end of the area. R_{er} is used to evaluate the passing performance of a total space when lines of sight pass through it. The number of reflections is m . Strength $F(i)$ is $F(i) = am$. Formula (3) shows the resistance value R_{er} .

$$R_{er} = \frac{N}{\sum_{i=1}^N F(i)} \quad (3)$$

In this research, we use the resistance value in the case that the line of sight is not reflected.

2.5 Simulation system and experimental method

To measure the resistance values and visualize the passing performance when self-driven particles or lines of sight pass through a space, we developed a system consisting of an open source programming language and an integrated development environment called Processing. In this study, as a first step of measurement, we created the area space. The area space is a plane 16 m×12 m, as shown in Figure 2.

Self-driven particles are arranged at a constant interval on the right edge and move to the left. We change the number of self-driven parti-

cles and measure the resistance value, and then we examine the number of particles necessary for measurement of the resistance value.

In the same way, lines of sight are arranged at a constant interval on the right edge and move to the left. We change the number of the lines of sight and measure the resistance value, and we again examine the number necessary for measurement of the resistance value.

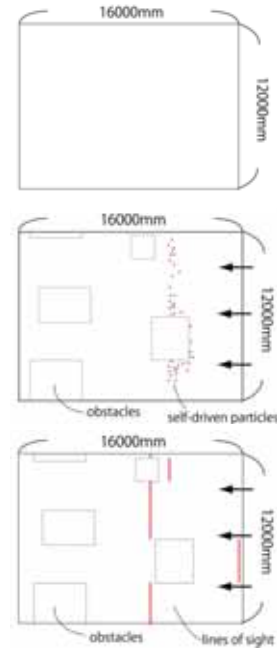


Figure 2: Area of the space and visualization system of self-driven particles and lines of sight.

2.6 Case study

As a case study, we make arrangement patterns that have some obstacles. The patterns are shown in Figure 3.

We make self-driven particles or lines of sight that pass through the area, then we measure the resistance values and visualize the passing performance.

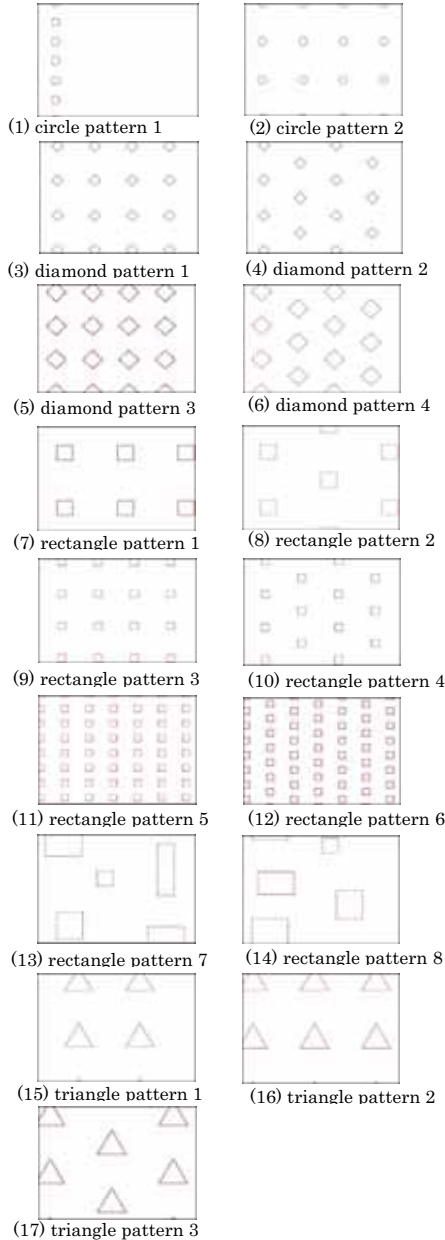


Figure 3: Arrangement patterns used for the case study

3. RESULTS AND DISCUSSIONS

3.1 Validating the number of passing objects required for measurement

We verified the number of self-driven particles and lines of sight required to measure the resistance values. First, we measured the resistance values of arrangement pattern (17) while changing the number of self-driven particles and lines of sight. Figures 4 and 5 show the results. In the case of self-driven particles and lines of sight, as the number increases, the rate of change of the resistance value decreases. In this study, we use a number of the particles and the lines of sight that is near convergence and a number that is relatively small. The number of self-driven particles is 40 and the number of lines of sight is 300.

Self-driven particles move in response to obstacles and towards the average heading of their local flockmates by the rule of alignment. If each starting point of self-driven particles slightly changes, the change of the resistance value is small. Furthermore, the resistance value when self-driven particles pass through is calculated from the arithmetic mean of the average speed of all self-driven particles. Several changes of self-driven particles have little effect on the resistance value. For this reason, by using a few self-driven particles (40 particles), we can measure a value close to the correct resistance value.

On the other hand, because the lines of sight are straight, a change of each starting point of a line of sight greatly influences the resistance value. If the number of lines of sight is small, each interval between the lines of sight is large and the percentage of lines of sight blocked by the obstacles can change. The resistance value of the lines of sight is calculated from the percentage of lines that pass through a space. If the total number of lines of sight is small, a change of the number of lines of sight greatly influences the resistance value. For this reason, to measure the exact value, we have to use many lines of sight.

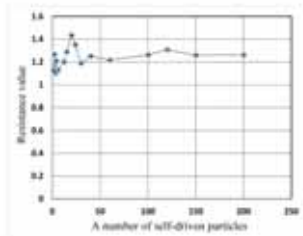


Figure 4: Scatter diagram of number of self-driven particles and resistance values

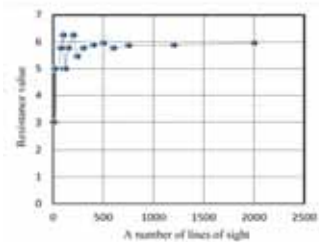


Figure 5: Scatter diagram of number of lines of sight and resistance values

3.2 Results of the case study

Figure 6 shows the resistance value of 17 arrangement patterns when self-driven particles or lines of sight pass through. Each number in Figure 6 is the number of one of the arrangement patterns in Figure 3. In this graph, we separate arrangement patterns into four groups and analyzed the characteristics of each group.

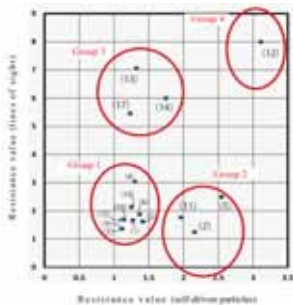


Figure 6: Scatter diagram of resistance values when self-driven particles and lines of sight pass through the area

Group 1 consists of arrangement patterns that have a low resistance value when self-driven particles or lines of sight pass through. It is easy for the self-driven particles and lines of sight to pass through the arrangement patterns. One example is arrangement pattern (7). Figure 7 shows the simulation screen of this pattern. The line in the figure is the locus of the self-driven particles or the lines of sight. In this case, the form of the obstacle is a rectangle. Because its sides are arranged at right angles, the self-driven particles hardly change direction or collide. Parts of the lines of sight are cut off. However, the parts between the obstacles are arranged in parallel lines. Once the self-driven particles and the lines of sight are interrupted, they pass through without further interruption.

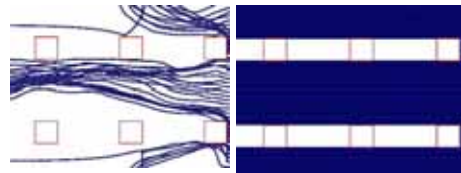


Figure 7: Simulation screen of arrangement pattern (7) rectangle pattern 1
(Left : self-driven particles, Right: lines of sight)

Group 2 consists of arrangement patterns that have a high resistance value when self-driven particles pass through and a low resistance value when lines of sight pass through. That is, it is difficult for self-driven particles to pass through and it is easy for lines of sight to pass through. One example is arrangement pattern (5). Figure 8 shows a simulation screen of this pattern. In this case, the form of the obstacle is a diamond. Because the diamonds are arranged at 45 degrees, the change of direction of the self-driven particles is smaller than those in

group 1. However, after a collision is prevented, the self-driven particles converge with each other, and their density increases and their speed decreases. As a result, it takes time for them to arrive at the objective point and the resistance value becomes high. Because the obstacles are parallel, once the lines of sight are interrupted, they pass through without interruption.

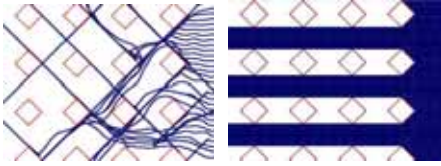


Figure 8: Simulation screen of arrangement pattern (5), diamond pattern 3

Group 3 consists of arrangement patterns that have a low resistance value when self-driven particles pass through and a high resistance value when lines of sight pass through. It is easy for self-driven particles to pass through and it is difficult for lines of sight to pass through. An example is arrangement pattern (14). Figure 9 shows the simulation screen of this pattern. The second example is arrangement pattern (17). Figure 10 shows the simulation screen of this pattern.

In the case of arrangement pattern (14), shown in Figure 9, the obstacles are rectangles. Because the sides of the rectangle are arranged at right angles, the self-driven particles hardly change direction or collide. However, because the density of the obstacles is low, the number of times that the self-driven particles are interrupted is small. Therefore, the self-driven particles arrive at the objective point in a relatively short time. On the other hand, many of the lines of sight are blocked because the obstacles are alternately arranged.

In the case of arrangement pattern (17),

shown in Figure 10, the obstacles are triangles. The change of direction of the self-driven particles is smaller than that for arrangement pattern (7) of group 1. Unlike arrangement pattern (5) of group 2, after a collision is prevented, the self-driven particles meet only a few other particles, and their speed does not decrease. As a result, the particles arrive at the objective point in a relatively short time. On the other hand, many lines of sight are blocked because the obstacles are alternately arranged.

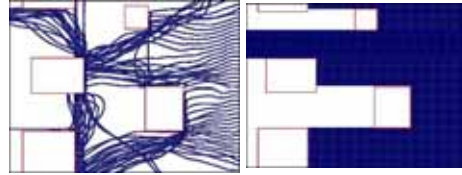


Figure 9: Simulation screen of arrangement pattern (14), rectangle pattern 8

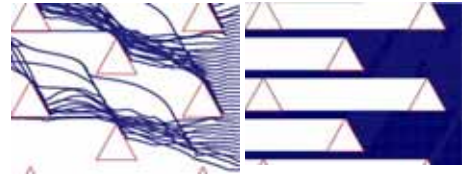


Figure 10: Simulation screen of arrangement pattern (17), triangle pattern 3

Group 4 consists of arrangement patterns that have a high resistance value when self-driven particles or lines of sight pass through. It is difficult for self-driven particles and lines of sight to pass through these arrangement patterns. The only example in this paper is arrangement pattern (12). Figure 11 shows a simulation screen of this pattern. In this case, the form of the obstacle is a rectangle. Because the sides are arranged at right angles, self-driven particles hardly change direction or collide. In comparison with the other forms, because the obstacles are arranged alternately

and with high density, the self-driven particles and lines of sight are greatly interrupted.

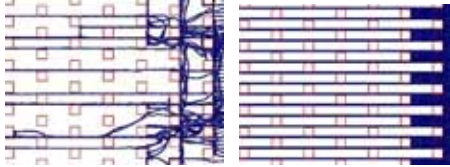


Figure 11: Simulation screen of arrangement pattern (12), rectangle pattern 6

As a result of this case study, the resistance value when self-driven particles pass through is affected by the size, the form, and the arrangement of obstacles, and the resistance value when the lines of sight pass through is affected by the size and arrangement of obstacles.

4. CONCLUSION

By conducting a case study, we obtained the following results.

- (1) If self-driven particles are interrupted, the resistance value becomes large. The factors that interrupt self-driven particles are size, form, and arrangement of obstacles.
- (2) If lines of sight are blocked, the resistance value increases. The factors that block the lines of sight are size and arrangement of obstacles.
- (3) Factors that affect the resistance value of self-driven particles are different from factors that affect the resistance value of the lines of sight.
- (4) One of the factors that makes the resistance value of self-driven particles larger than the resistance value of lines of sight is as follows.
When self-driven particles change movement direction to avoid collision with obstacles and the density of the self-driven particles increase, the speed of the self-driven particles decreases.
- (5) One of the factors that makes the resistance

value of lines of sight larger than the resistance value of self-driven particles is as follows.

If the density of the obstacles is low and the obstacles are alternately arranged, the lines of sight are blocked.

The purpose of this study was to propose an evaluation index of passing performance and to develop a method of evaluation. We defined resistance value of self-driven particles and lines of sight as an index of passing performance. We developed the system for measuring resistance value and visualizing the passing performance. As a case study, we measured resistance value of some arrangement patterns. From those result, we analyzed features of resistance value of self-driven particles and lines of sight.

We next show some issues that need to be addressed. First, we have to conduct measurements in three-dimensional space because passing is a phenomenon in three-dimensional space. Second, we have to model the behavior of people or cars as self-driven particles. Using this model, we can calculate the resistance values of some real spaces and compare these resistance values with the analysis of real space.

REFERENCES

- [1] T. Koizumi, "The prototypes of natural light which generates unity or direction to an architectural space", NO. 482 Journal of Architecture and Planning, 4, 95-103, 1996.
- [2] A. Coleman, "Utopia on Trial: Vision and Reality in Planned Housing", Hilary Shipman Ltd, 1985.
- [3] S. Yamamoto, "Study on compact city from a viewpoint of permeability as spatial characteristic - 1 - Definition of the concept of permeability", Summaries of technical papers of annual meeting of the Architectural Institute of Japan, 917-918, 2005.
- [4] K. Nishinari, "Collective dynamics of self-driven particles and forming a jam", Publications of the Research Institute for

Mathematical Sciences, vol. 1472, 118-128, 2006.

- [5] C. W. Reynolds, “*Flocks, herds, and schools: A distributed behavioral model*”, Computer Graphics (Proc. SIGGRAPH), 25-34, 1987.

ABOUT THE AUTHORS

1. Takuya Matsumoto, “B.Eng.”, is a student at Osaka University, 3-6-8 Makiochi, Mino, Osaka 512-0004, Japan
2. Kensuke Yasufuku, “D.Eng.”, is an assistant professor at Osaka University, Cybermedia Center, 5-1 Mihogaoka, Ibaraki, Osaka 567-0047, Japan
3. Hirokazu Abe, “D.Eng.”, is a professor at Osaka University, Cybermedia Center, 5-1 Mihogaoka, Ibaraki, Osaka 567-0047, Japan

EVALUATION OF SPATIAL ABILITY BY USING A SILHOUETTE OF A SOLID FIGURE IN GRAPHIC SCIENCE EDUCATION

Akira TAKAHASHI¹ and Hirokazu ABE²

¹Kyoto Center for Community Collaboration, Japan ²Osaka University, Japan

ABSTRACT: It has frequently been said that one of the significant objectives of graphic science education is the cultivation of spatial ability. We note that the importance of this ability increases with the spread of 3DCG/CAD technology. The Mental Cutting Test (MCT) is commonly used to assess spatial ability. This is an objective test, in which the subject is asked to imagine that a test solid is cut along an indicated plane; the subject is then asked to choose the correct cross-section from among five alternatives. There have been many studies of this test. We developed the Mental Silhouette Test (MST) to explore spatial ability from a different point of view. The MST is also an objective test, and the subject is asked to imagine the silhouette of some test solids that are used in the MCT; the subject is then asked to choose the incorrect silhouette from among five possible orthographic projections. The MST and MCT were administered to 120 students taking a graphic science course at a university in 2012, and the data were then analyzed according to item response theory (IRT). The findings are as follows. 1) The correlation coefficient between the MST and MCT test scores was 0.59. However, the correlation coefficient between the MST scores and those on the final exam in the graphic science class was 0.25. 2) The IRT analysis showed that the difficulty (as measured by the threshold parameter) of the MST was higher than that of the MCT, but the discrimination (as measured by the slope parameter) of the MST was lower than that of the MCT.

Keywords: Spatial Ability, MCT, Graphic Education

1. INTRODUCTION

It has frequently been said that one of the significant objectives of graphic science education is the cultivation of spatial ability [1]. We note that this importance increases with the spread of 3DCG/CAD technology. In graphic science education, the Mental Cutting Test (MCT) is commonly used to assess spatial ability. This is an objective test in which the subjects are asked to imagine that a test solid is cut along an indicated plane, and they are then asked to choose the correct cross-section from among five alternatives (Fig. 1). There have been many studies of the MCT [2], [3].

We developed an objective test that uses silhouettes to explore spatial ability from a different point of view. A silhouette is the projection of a

three-dimensional object onto a two-dimensional screen. Since the silhouette is a single shade of black, it is difficult to imagine the shape of the object. For instance, consider hand shadows, which create images of animals; it is difficult to imagine the shape of the hand from the silhouette (Fig. 2).

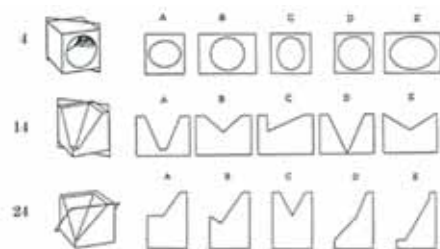


Figure 1: Examples from the MCT.

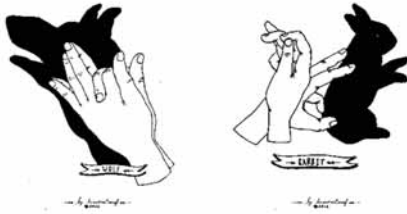


Figure 2: Examples of hand shadows [4].

People tend to interpret what they see based on the simplest form that is already known. It is assumed that this is one of the factors in an illusion. There is a possibility that, for the students in the graphic science class, the optical illusion effect obstructs spatial recognition.

The objective test we developed is the Mental Silhouette Test (MST). In this test, the subjects are asked to imagine the silhouettes of some of the test solids used in the MCT, and they are then asked to choose the incorrect silhouette from among five possible orthographic projections. This paper explores spatial ability in a graphic science education setting by using item response theory (IRT) to analyze the relationship between the MCT and the MST. IRT is used to analyze other equivalency tests, such as TOEFL (Test of English as a Foreign Language) [5], and it analyzes the difficulty of each test item by determining the rate of correct answers rate and uses this to form a characteristic value. In a previous study, Shiina [6] used IRT to estimate a parameter for each item in the MCT.

2. METHOD

Figure 3 shows an example of the MST. It consists of 20 questions (items), each of which ask that the incorrect silhouette be chosen from among five alternatives (silhouettes) that are projections of the test solid rotated by an arbitrary angle. The silhouettes are the orthographic projections.

Table 1 shows the number of subjects in this study. The MST and MCT were administered to freshmen taking a graphic science course at Osaka University in 2012. The MCT was administered in April 2012 at the beginning of the term,

and the MST was administered with the final exam in February 2013. The subjects of this research were 120 students who took both the MCT and the MST.

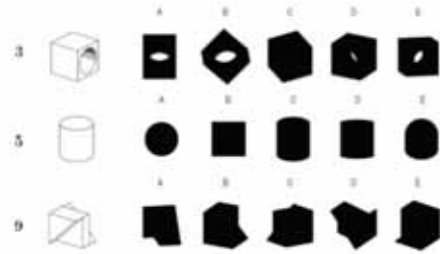


Figure 3: Examples from the MST.

Table 1: Number of subjects

Name of Tests	The date of implementation	number of subject	(person)
			valid subject
MST	5, Feb. 2013	120	120
MCT	10, Apr. 2012	126	120
Terminal exam of the class	5, Feb. 2013	120	120

2.1 Item Response Theory (IRT)

IRT analyzes the features of each item on the test and replaces the probability of the correct response to each item with a logistic function [7]. In this study, we adopted a three-parameter logistic model based on the probability that a correct response corresponds to the student's ability. In a three-parameter logistic model, the ability of the student is denoted by θ , and the probability of a correct response to item j is represented by $P_j(\theta)$, as follows:

$$P_j(\theta) = c + \frac{1 - c}{1 + \exp(-Da_j(\theta - b_j))}$$

$$-\infty < \theta < \infty \quad \text{----} (1)$$

where a_j and b_j represent the characteristics of each item and c represents chance, as described in more detail in the following. The parameter a_j represents the degree to which item j discrimi-

nates between individuals. This parameter characterizes the maximum slope of the item response function (IRF). The parameter b_j represents the difficulty of the item. It is the value of θ at which the slope of the IRF is a maximum. For multiple-choice items, the parameter c represents the likelihood of getting a correct answer by randomly choosing an answer. In the case of a five-option multiple-choice item, there is a 1/5 chance that a candidate could guess the correct answer, so c would be approximately 0.2. Here, D is a constant equal to 1.7, which aligns the scale of the ability value θ with the scale of a standard cumulative model.

In this study, we used a three-parameter logistic model, and we estimated the parameters a (slope) and b (threshold) by using the application BILOG version 3.11 [8]. However, it was necessary to presort values if the parameter was too high or too low, as follows:

- 1) slope < 0.2
- 2) threshold < -3.0 or threshold > 3.0,

3. RESULTS AND DISCUSSION

3.1 Scores of the MCT and MST

Table 2 shows the scores of the MCT and MST. For the MST, the average score was 57.9%, the highest score was 85%, and the lowest score was 15%. For the MCT, the average score was 75.1%, the highest score was 100%, and the lowest score was 24 %.

Table 2: Scores of the MCT and MST

Name of Tests	valid subject	(converted full score 100)	
		Average Score	Standard Deviation
MST	120	57.9	13.3
MCT	120	75.1	16.5
Terminal exam of the class	120	56.5	22.0

Table 3: Correlation between scores on the MCT and those on the MST

correlation coefficient	r	0.59
number of samples	n	120
significance level	α	0.05
amount of freedom	ϕ	118
t-statistic	t	8.014
t-ratio	t	1.98
probability both sides	p	8.940E-13 < 0.05
confidence interval 95%		
0.46 < σ < 0.69		

Table 3 shows the correlation between the scores on the MCT and those on the MST. The correlation coefficient (r) is 0.59. The 95% confidence interval of correlation coefficient (σ) is between 0.46 (lower bound) and 0.69 (upper limit). The probability of no correlation (p) is 8.940E-13. This suggests that there is a correlation between the scores on these two tests.

3.2 Error Analysis for the MST

Figure 4 shows the rate of correct answers for each item in the MST.

Table 4: Number of respondents of 5 choices in each item

Item Number option	(number of subject)																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	8	0	26	33	0	15	57	0	18	25	104	2	75	37	34	98	56	0	7	74
B	3	21	12	6	2	44	38	0	24	14	6	10	3	76	43	15	57	2	42	24
C	11	2	42	0	2	30	3	105	29	22	1	2	41	2		1	2	5	27	11
D	85	9	14	70	2	3	8	0	13	17	6	101	1	2	25	6	4	93	9	6
E	13	88	26	11	113	28	14	14	34	42	2	5	0	1	18	0	0	20	31	5
Z	0	0	0	0	1	0	0	1	2	0	1	0	0	2	0	0	1	0	4	0

shows the correct answer choice
 shows the first answer choice

There is a large gap between the rate for question No. 5 (94.2%) and that of No. 15 (20.8%). Eight items (Nos. 3, 6, 7, 9, 10, 15, 17, and 19) have rates that are less than average.

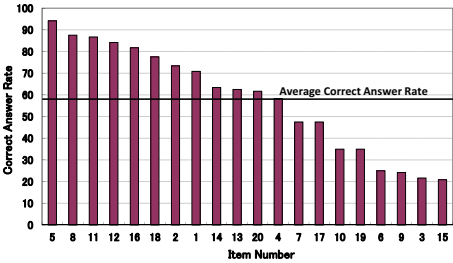


Figure 4: Rate of correct answers for each item in the MST.

Table 4 shows the number of respondents for each of the five choices for each item. The black cells show the correct answer, and the gray cells show the choice selected by the largest number of students. In four of the items, the number of students who selected wrong answers was larger than those who selected the correct answer. This is shown in Fig. 5.

For item No. 15, 77 students chose options A or B, both of which were wrong. It was apparently difficult for the students to imagine the original shape and difficult to find a vertex of the solid, because these silhouettes are different from the original shapes, as was seen with the hand shadows (Fig. 2). A rotation of the test solid may appear to produce a mistakable specialized silhouette. There is a possibility that the illusion effect obstructs the correct answer (Fig. 5). To obtain the correct answer, we must find the straight line that connects point *a* with point *b*; this corresponds to a ridge line on the test solid, as seen in the axonometric view of Fig. 5.

Regarding item No. 3, we note that 42 students choose C, one of the wrong options. This seems to be an intuitive answer, since we can see that the figure has a hole. However, the shape of the wrong choice is a little different from the original figure (see Fig. 4)

We can find the answer by comparing the size of

the holes with the arc in the silhouette, such as is seen in the axonometric view of Fig. 3.

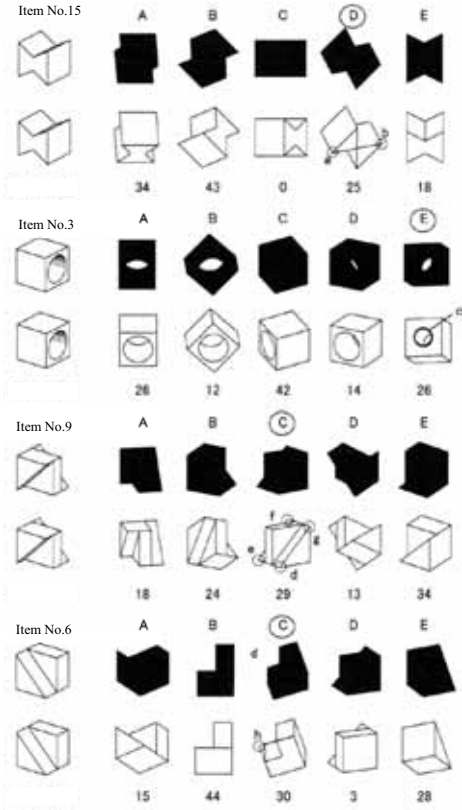


Figure 5: Four items for which the number of wrong answers was larger than the number of correct answers

Regarding item No. 9, we note that each choice was selected by many students. Because the silhouette from an arbitrary angle is complex, it is difficult to choose from the alternatives. For this item, the correct answer can be obtained by finding the position of the vertex *d* of the solid. This requires a quantitative judgment in comparing the lengths of lines *ef* and *fg*.

Regarding item No. 6, 44 students choose B, one of the wrong options. This option seems correct

because it is a figure seen from the direction of the ridge line of the triangular prism. The correct answer can be obtained by finding the position of the vertex h of the solid. This requires a quantitative judgment.

The items in the MST for which there were many wrong answers suggest that it is difficult to solve some questions intuitively without a quantitative judgment. There is also the possibility that the illusion effect of the silhouette obstructed the correct answer, since the rotation of the test solid might appear to be a mistakable specialized form of the silhouette.

3.3 Parameters of the Items of the MST

Table 4 shows the score of each parameter (slope and threshold) for each item, as estimated by IRT.

Table 5: Scores of slope and threshold for each item (MST)

Item Number	Correct Answer Rate	SLOPE	THRESHOLD
1	0.71	0.403	-0.770
2	0.73	0.352	-1.109
3	0.22	1.069	1.725
4	0.58	0.437	0.297
5	0.94	1.118	-1.980
6	0.25	1.056	1.214
7	0.48	0.746	0.628
8	0.88	0.730	-1.704
9	0.24	1.060	2.181
10	0.35	0.758	1.519
11	0.87	0.666	-1.748
12	0.84	0.365	-2.332
13	0.83	0.783	-0.055
14	0.83	0.878	-0.203
15	0.21	-	-
16	0.82	0.900	-1.060
17	0.48	1.029	0.528
18	0.78	0.550	-1.117
19	0.35	0.905	2.561
20	0.82	0.957	-0.061

■ : High score ■ : Low score

Figure 6 shows the IRF for each item. The horizontal axis presents the ability of each student as a scale score, and the vertical axis shows the probability of a correct response to each item (PROB). (Note that item No. 15 was excluded and values were recalculated, because the estimated slope and threshold were out of the preset range.)

In Nos. 3, 5, and 9 (Fig. 7), the slope (discrimination) was high. In Nos. 3 and 9, the value of the threshold (difficulty) was also high.

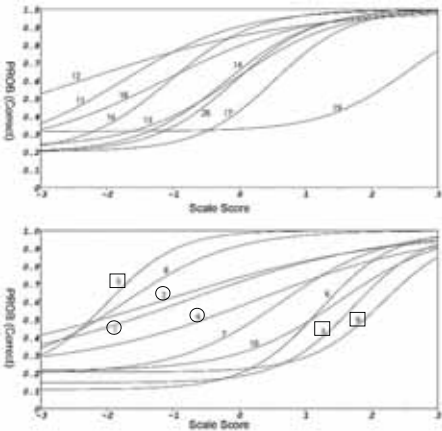


Figure 6: IRF for each item in the MST.

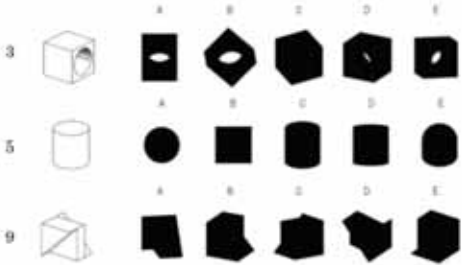


Figure 7: The items with higher slope scores.

In Nos. 3, 5, and 9, the curve shows rapid changes in the value of PROB (Fig. 6). Item No. 5 changed at the point where the scale score was lower than that of the other items. This means that No. 3 may be a better item measure the students of low ability.

In Nos. 1, 2, and 4 (Fig. 8), the value of the slope (discrimination) is low. The value of the threshold (difficulty) is medium.

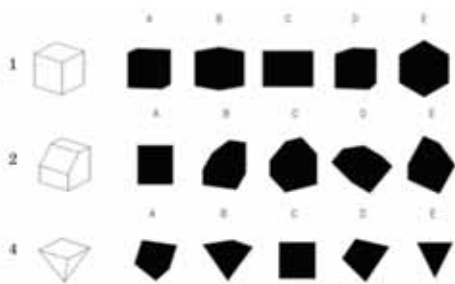


Figure 8: Items with lower slope scores.

In Nos. 1, 2, and 4, the shape of the curve shows that the value of PROB gradually changes (Fig. 6). This means that this item does not work well to discriminate between students of high or low ability.

3.4 Parameters of the Items of the MCT

Table 6 shows the scores of the parameters (slope and threshold) of the MCT, as estimated by IRT. Figure 9 shows the IRF of the MCT. In Nos. 17, 20, and 21 (Fig. 10), the slope (discrimination) is high, and the threshold (difficulty) is low. In Nos. 17, 20, and 21, the shape of the curve shows that the value of PROB changes suddenly and at the same scale score (Fig. 9). However, in Nos. 4, 14, and 24 (Fig. 11), the slope (discrimination) is low.

In Nos. 4 and 14, the threshold (difficulty) is low, but in No. 24, it is high. In all three (4, 14, and 24), however, the shape of the curve indicates that the value of PROB does not change suddenly (Fig. 9).

Items No. 9 of the MST and No. 20 of the MCT are the same solids, and they have high slope (discrimination) scores (Fig. 12). This solid has the form of a cube divided into three parts and cut diagonally from different directions. It was good for distinguishing between students.

Items No. 3 of the MST and No. 4 of the MCT are also the same solids. However the slope of No. 3 (MST) is high while that of No. 4 (MCT) is low.

Table 6: Scores of slope and threshold in each item (MCT)

Item Number	Correct Answer Rate	SLOPE	THRESHOLD
1	0.93	0.859	-2.431
2	0.68	0.737	-0.385
3	0.98	1.031	-2.854
4	0.87	0.613	-1.722
5	0.88	0.927	-1.477
6	0.93	0.788	-2.243
7	0.90	0.956	-1.555
8	0.91	0.808	-1.813
9	0.54	0.730	0.477
10	0.92	0.733	-2.144
11	0.82	0.743	-1.255
12	0.74	0.872	-0.639
13	0.54	0.949	0.189
14	0.69	0.647	-0.433
15	0.84	0.882	-1.415
16	0.70	1.127	-0.303
17	0.81	1.335	-0.875
18	0.83	0.747	-1.252
19	0.87	1.281	-0.285
20	0.85	1.419	-0.988
21	0.73	1.485	-0.294
22	0.73	1.011	-0.572
23	0.34	1.159	0.973
24	0.51	0.595	0.585
25	0.43	1.263	0.469

■ : High score ■ : Low score

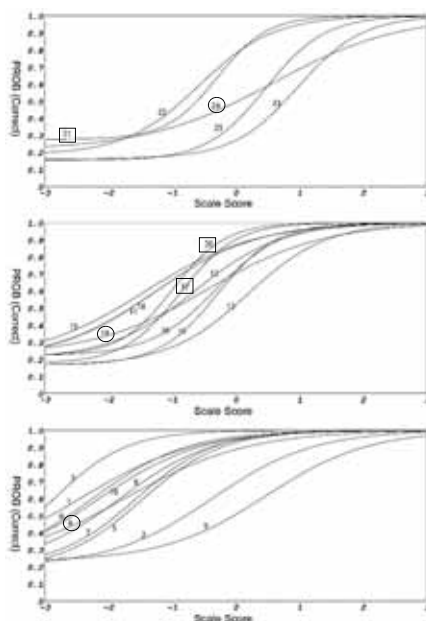


Figure 9: IRF for each item of the MCT.

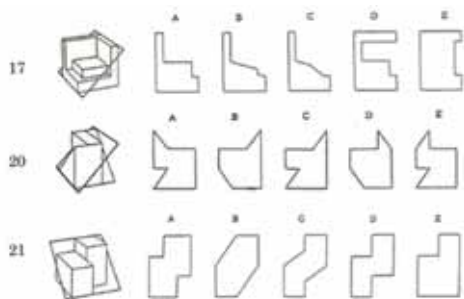


Figure 10: Items with higher slope scores.

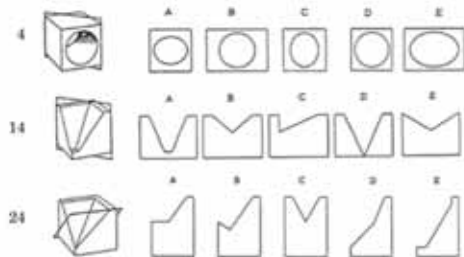


Figure 11: Items with lower slope scores.

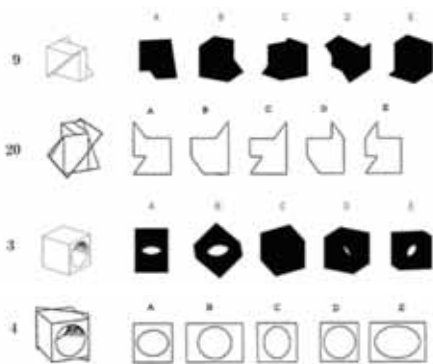


Figure 12: Test solid that is common between the MCT and the MST.

The latter requires a qualitative estimation about the location of a hole, while the former requires a quantitative estimation about the size of the hole. Thus, the characteristics of these items are different.

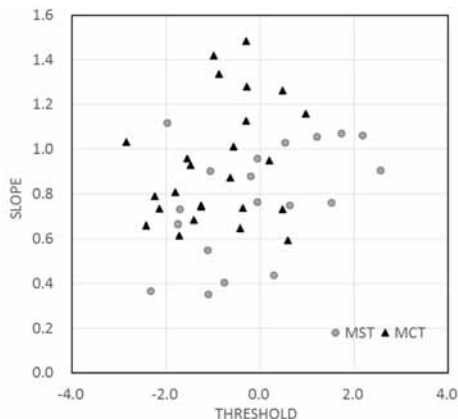


Figure 13: Distribution of the item parameters in the MCT and MST.

Figure 13 shows the distribution of the item parameters in the MCT and MST. For the MST, the slope is distributed within the range of 0.35 to 1.11, and the threshold is distributed within the range of -2.33 to 2.56. For the MCT, the slope is distributed within the range of 0.59 to 1.48, and the threshold is distributed within the range of -2.85 to 0.97.

Thus, the MST discrimination score is lower than that of the MCT, and the MST difficulty score is higher than that of the MCT.

3.5 Graphic Science Class and the MST

At Osaka University, the aim of the graphic science class is to cultivate the spatial visualization ability, based on descriptive geometry. It covers the theory of orthogonal projections, military projections, cavalier projections, and axonometric and perspective drawing using computer-aided design tools.

Table 7 shows the results of a correlation analysis between the scores of the final examination in the graphic science class and the MST.

The correlation coefficient (r) between the final examination and the MST is 0.25. The 95% confidence interval of the correlation coefficient (σ) is between 0.07 (lower bound) and 0.41 (upper

limit). This suggests that there is little correlation between the final examination of the graphic science class and the MST.

Table 7: Correlation analysis between the final exam in the graphic science class and the MST

correlation coefficient	r	0.25
number of samples	n	120
significance level	α	0.05
amount of freedom	ϕ	118
t-statistic	t	2.833310633
t-ratio	t	1.980272249
probability both sides	p	0.0054 < 0.05
confidence interval 95%		
0.07 < σ < 0.41		

4. CONCLUSIONS

In this paper, the Mental Silhouette Test (MST) was developed to explore the spatial abilities of students. The MST and the MCT were both administered to 120 students, and the data were analyzed. The results are as follows.

1. The average scores of the MST and MCT are 57.9% (SD: 13.3) and 75.1% (SD: 16.5), respectively.
2. The correlation coefficient for the test scores of the MST and MCT is 0.59. However, the correlation coefficient for the MST and the final exam of the graphic science class was 0.25.
3. The items in the MST for which there were many wrong answers were those which were difficult to solve intuitively without a quantitative judgment. There is also the possibility that the illusion effect obstructed the correct answer, since the rotation of the test solid might appear to be a mistakable specialized form of the silhouette.
4. IRT analysis showed that the difficulty parameter (threshold) of the MST is higher than that of the MCT, while the discrimination parameter (slope) of the MST is lower than that of the MCT.
5. We found some solids that were identical in the MCT and MST and for which IRT indicated that the discrimination parameter (slope) was high.

REFERENCES

- [1] K. Suzuki: "Nurturing a Spatial Ability by the Graphic Science Education". Proceedings of Annual meeting of JSGS, pp21-26, 2000.
- [2] K. Suzuki: "Evaluation of Students' Spatial Abilities by a Mental Cutting Test - Review on the Surveys in the Past Decade-", The 11th International Conference on Geometry and Graphics, (Invited lecture), Guangzhou, China, August, 2004.
- [3] E. Tsutsumi, W. Ishikawa, H. Sakuta and K. Suzuki: "Analysis of Causes of Errors in the Mental Cutting Test -Effects of View Rotation-". Journal for Geometry and Graphics, Vol.12, No.1. 109-120, 2008.
- [4] Lara Mendes, "Hand Shadows Creates Critics with Shadows" <http://www.trendhunter.com/trends/lara-mendes-hand-shadows> (25 May 2014)
- [5] James E. Carlson: "ETS R&D Scientific and Policy Contributions Series" ETS research report, No.RR-13-28, 2013.
- [6] K. Shiina, Application of Item Response Theory to Spatial Test, Proceedings of annual meeting of JSGS, pp63-69, 2001
- [7] Sukeyori Siba, "Item Response Theory", University of Tokyo Press, 1991
- [8] Robert J. Mislevy, R. Danell Bock, BI-LOG Version 3.11, Item Analysis and Test Scoring with Binary Logistic Models, 1997

ABOUT THE AUTHORS

1. Akira TAKAHASHI, Dr. Eng., is a coordinator at the Kyoto Center for Community Collaboration, 83-1 Umeminato-cho, Nishi Kiyamachi-dori, Kaminokuchi agaru, Shimogyo-ku, Kyoto, 600-8127, Japan
2. Hirokazu ABE, Dr. Eng., is a professor at Osaka University, Cyber Media Center, 5-1 Mihogaoka, Ibaraki, Osaka 567-0047, Japan

EVALUATION ON TRIANGLE-TRIANGLE INTERSECTION TESTS ALGORITHMS

Haiyan Yu¹ Yuanjun He² Yongxing Wang¹

¹Donghua University, China ²Shanghai Jiaotong University, China

ABSTRACT: Over the years, as the development of the algorithms of the triangle-triangle intersection tests, there comes out lots of complex, abstract and high-speed algorithms. However, there is no general standard for comparing these algorithms' stability. This paper constitutes a visual method to evaluate objectively. For the stability test, we pay more attention on the special singularity. Based on the projection theories, this paper brings out a new kind of classification. The first classification clue is the different positional relationship of planes of the triangle pairs (parallel, coplanar, intersect). Then a computing coordinate which is corresponding to a three-plane projection system of V-H-W is constructed and a plane has seven types of positions to three projecting planes V-H-W. One triangle is fixed and changes the other triangle to get various projecting types. Particularly, once two planes of a triangle pair is selected; one triangle will be moved to the other to get various critical positions such as point-point collision, point-line collision, line-line collision and point-plane collision. The size of one triangle is also changed to get some special triangle pairs that are prone to be error in a triangle-triangle intersection test algorithm. It provides an opportunity for the researcher to self-check his/her algorithm before publishing one with loopholes, as well as provide a solution for the testers in company to reduce their workload of the detection, and improve the efficiency.

Keywords: collision detection, 3D triangles, intersection test, efficiency, stability, evaluation.

1. INTRODUCTION

A virtual scene usually consists of a certain amount of static environmental objects as well as moving active objects, each of which is composed by thousands of basic geometric elements. As most geometric elements consist of triangular meshes, it is necessary for many collision detection algorithms to include a triangle-triangle intersection test. Consequently, an efficient triangle-triangle intersection test will ultimately contribute to collision detection and has application in CAD, the finite element technique, 3D games and animations.

From the primitive brute-force to classic Möller and Held algorithm and then algorithms proposed by Roy, Shen, Guigue, Troppo, the research of triangle-triangle intersection test is always a concern [1-6]. As the development of related application, many new algorithms are

proposed to enhance the detection velocity. Unfortunately, the algorithms are more complex and often expressed by complicated numerical functions without geometrical meaning. Consequently, it is somehow difficult to detect the stableness from the view of theory. On the other side, it is not evident from the literature that published intersection tests deal with triangle pairs that are in geometric singularities such as in 'grazing' contact.

Based on this situation, this paper will proposed a visual method to classify varied triangle pairs based on projection. In our method, position relation of a three dimensional triangle pair is reduced to a planar segment pair or between a segment and a triangle. Each 3D triangle pair could be drawn easily and visually on the plane. Particularly, all singularities could be recognized and graphically represented by illustrative examples. Further-

more, a test strategy is provided including a complete test data workflow regarding the staleness and speed to triangle-triangle intersection algorithms.

2. RELATED WORKS

Due to massive complicated computer models in engineer, one error in triangle intersection tests may cause incalculable loss of the entire project.

Steven Robbins and Sue Whitesides note that implementors of triangle/triangle intersection tests often opt to forego exact calculations for speed reasons. It is widely known that such code will fail for certain inputs [7]. Christer Ericson also notes that frequently they test the primitives in some large number of random configurations and report the average time it took to do a test. This kind of testing is very unlikely to hit the cases where the tests fail due to robustness issues [8].

According to the published literature, in order to test the actual performance of a new algorithm, the general test method is to take a set of triangle pairs as test data and select one or more typical algorithms as comparison. But currently, the sample for experiments is usually designed by the one who develop a new algorithm and there is not an objective sample.

3. PROJECTION THEROYAND THEOD

In order to describe positions of a triangle pair completely and visually, a projection reduction method is proposed to classify and interpret it. The relation of geometries is irrelevant to the coordinate system. But a proper coordinate system can simplify its geometric representation and algebra computing. A computing coordinate system is constructed corresponding to a three projection system.

The computing coordinate system is set as: to any triangle pair of A and B, the plane A is taken as coordinate plane xy, and the normal vector of plane A is take as axis Z.

Figure 1 shows the computing coordinate system and the corresponding three projections, where coordinate plane xy is corresponding to projection plane H, plane zx to V, plane yz to

W, and A_H for triangle A 's projection on plane H.

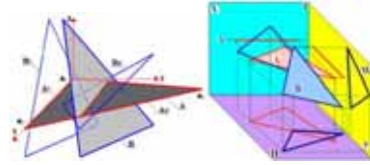


Figure 1: the computing coordinate system and its projection system.

According to descriptive geometry, a plane may have seven position relations to a three dimension systems. They are respectively horizontal plane (a plane parallel to H), frontal plane (a plane parallel to V), profile plane (a plane parallel to W), H-perpendicular plane (a plane perpendicular only to H), V-perpendicular plane (a plane perpendicular only to V), W-perpendicular plane (a plane perpendicular only to W) and general-plane (a plane that is neither parallel nor perpendicular to the three projection planes).

Based on the above concept, in the projection system, triangle A is fixed to a horizontal plane. Triangle B is changed to get seven projection positions. Thus, we get seven relations to this triangle pair. The projections of triangle A on plane V and W are both segments, while on plane H its projection is just the triangle itself. The projections of triangle B to each projection plane is a triangle or a segment. Table 1 shows the projection relations of a 3D triangle pair, where L stands for a segment and \triangle stands for a triangle, P_A stands for the plane of triangle A and P_B the plane of triangle B.

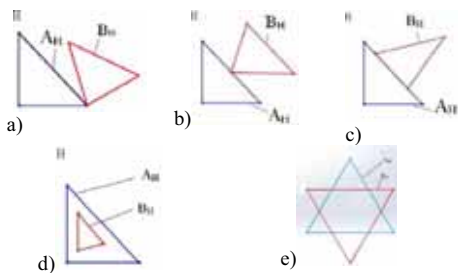
4. DESIGN OF TESTING SAMPLES

A good classification will contribute to the completeness of samples. Our method is fix triangle A and by change triangle B to get all kinds of spatial relations. The spatial position of plane B is first changed, then on a particular plane, triangle B is moved and its size is also taken into consideration.

When plane A and plane B parallel, triangle A and B are not interest for sure. So we analyze the two plane are co-plane and intersecting for

detail. The move of triangle B is mapped by the projection.

When triangle A and B are on the same plane, triangle A and B either intersect or not. In the case they are intersecting, there are some critical relations as graphical represented in Figure 2. In our projection system, triangle A and B are on projection H. Consequently, their projections on H just maps spatial state.

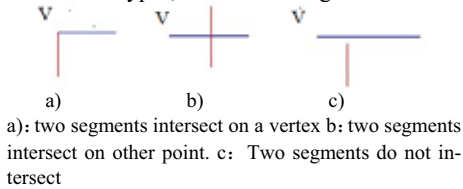


a): Point-point collision; b): Point-line collision; c): Line-line collision; d): a triangle inside the other; e): None vertex of a triangle is in the other, but two triangles still intersect.

Figure 2: Critical relations as two triangles are on the same plane

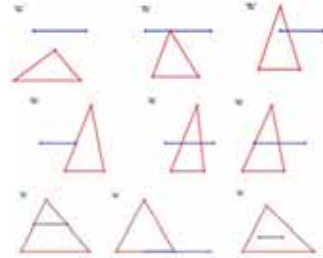
When two planes intersect, the projections of two triangles include six relation types from t-2 to t-7 as showed in table 1. We first move triangle B on the projection plane that the projection relation is L-L and L- Δ .

Take t-3 as an example, P_B is a profile plane, and on projection plane V, the projection is L-L and there are three sub-types shown as Figure 3. On projection plane W, the projection is classified to nine types, as shown in Figure 4.



a): two segments intersect on a vertex b): two segments intersect on other point. c): Two segments do not intersect

Figure 3: projection type of two triangles on V when triangle B is a profile plane



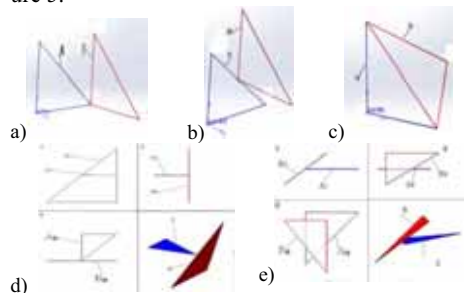
a): Point-point collision; b): Point-line collision; c): Figure 4: projection type of two triangles on W when triangle B is a profile plane

Based on the above classification method, the triangle pairs are classified to 105 types including 56 intersecting pairs and 49 non-intersecting pairs.

5. DISCUSSION AND CONCLUSION

This paper gives a new method to select testing samples based on a three-step classification, ie. the relations of two planes, the types of projections and projection relations in a view of relevant moving.

Emphasis is placed on varies critical positions including colliding and other special situations. Collision is taken as a special kind of intersection, including point-point, point-line, point-plane, and line-line collision. Figure 2 shows the case when two triangles are in the same plane. When the planes of the two triangles intersect, the following collision positions are listed into our testing sample as shown in Figure 5.



a):Point-point collision; b):Point-line collision; d):line-line collision; e):line-plane collision

Figure 5: Critical relations as two planes of triangles intersect

ACKNOWLEDGMENTS

This work is supported by the Fundamental Research Funds for the Central Universities (2232012D3-12).

REFERENCES

- [1] MOLLER T. A Fast Triangle-Triangle Intersection Test [J]. *Journal of Graphics Tools*, 1997, 2(2): 25-30.
- [2] Held, M. ERIT: A Collection of Efficient and Reliable Intersection Tests [J]. *Journal of Graphics Tools*, 1997, 2 (4): 25—44.
- [3] TROPP O, TAL A, SHIMSHONI I. A Fast Triangle to Triangle Intersection test for collision detection [J]. *Computer Animation and Virtual Worlds*, 2006, 17(5): 527-535.
- [4] DEVILLERS O, GUIGUE P. Faster triangle-triangle intersection tests, *TR4488[R]. [S.l.]: INRIA*, 2002.
- [5] GUIGUE P, DEVILLERS O. Fast and Robust Triangle-Triangle Overlap Test Using Orientation Predicates [J]. *Journal of Graphics Tools*, 2003, 8(1): 25-42.
- [6] HEN Hao, HENG P A, TANG Z. A Fast Triangle-Triangle Overlap Test Using Signed Distances [J]. *Journal of Graphics Tools*, 2003, 8(1): 3-15.
- [7] Steven Robbins, Sue Whitesides. On the Reliability of Triangle Intersection in 3D. *ICCSA 2003*, LNCS 2669, pp. 923-930, 2003
- [8] Christer Ericson. Triangle-triangle tests, plus the art of benchmarking <http://realtimecollisiondetection.net/blog/?p=29>

ABOUT THE AUTHORS

1. Haiyan YU, Ph.D, associate professor. Main research field is CAD/Computer Graphics. yuhy@dhu.edu.cn
2. Yuanjun HE, professor. Main research field is CAD/Computer Graphics. yjhe@sjtu.edu.cn
3. Yongxing Wang Ph.D, associate professor. Main research field is CAD/Textile machine design.

Table 1: the projection relations of $\triangle A$ and $\triangle B$

Relation type	P_B to projection planes	The projection of $\triangle B$			The projection of $\triangle A-\triangle B$			Relations of P_A and P_B
		V	W	H	V	W	H	
t-1	horizontal plane	L	L	\triangle	L-L	L-L	$\triangle-\triangle$	Horizontal or co-plane
t-2	V-perpendicular plane	L	\triangle	\triangle	L-L	L- \triangle	$\triangle-\triangle$	intersecting
t-3	profile plane	L	\triangle	L	L-L	L- \triangle	$\triangle-L$	intersecting
t-4	W-perpendicular plane	\triangle	L	\triangle	L- \triangle	L-L	$\triangle-\triangle$	intersecting
t-5	frontal plane	\triangle	L	L	L- \triangle	L-L	$\triangle-L$	intersecting
t-6	H-perpendicular plane	\triangle	\triangle	L	L- \triangle	L- \triangle	$\triangle-L$	intersecting
t-7	general-plane	\triangle	\triangle	\triangle	L- \triangle	L- \triangle	$\triangle-\triangle$	intersecting

EXISTENCE AND UNIQUENESS OF CENTERS OF REGULAR POLYGONS IN SOME SUBCLASSES OF IM-QUASIGROUPS

Stipe VIDAK

University of Zagreb, Croatia

ABSTRACT: The concept of an IM-quasigroup, an idempotent medial quasigroup, is defined as a quasigroup whose elements satisfy the identities of idempotency and mediality. Motivated by the example $C(q)$ some geometrical concepts can be defined in an IM-quasigroup. Many subclasses of IM-quasigroups have been defined and studied, such as GS-quasigroups, quadratical quasigroups, hexagonal quasigroups and pentagonal quasigroups. In this paper a special emphasis is put on hexagonal and pentagonal quasigroups. The notions of parallelogram, midpoint of a segment, regular triangle and regular hexagon are defined in hexagonal quasigroups. Some finite examples of hexagonal quasigroups are given and existence and uniqueness of the midpoint of a segment and of the centers of regular triangles and regular hexagons are studied. The notions of regular pentagon and regular decagon are defined in pentagonal quasigroups. Existence and uniqueness of their centers are studied in some finite examples of pentagonal quasigroups. The results are justified using characterization of these subclasses of IM-quasigroups via abelian groups which possess certain types of automorphisms.

Keywords: IM-quasigroup, hexagonal quasigroup, pentagonal quasigroup, midpoint, regular polygon

1. INTRODUCTION

A *quasigroup* (Q, \cdot) is a groupoid in which each of the equations $a \cdot x = b$ and $x \cdot a = b$ has a unique solution x for every $a, b \in Q$. We say that (Q, \cdot) is an *idempotent medial quasigroup*, or shorter *IM-quasigroup*, if we have $a \cdot a = a$ and $ab \cdot cd = ac \cdot bd$ for every $a, b, c, d \in Q$.

The notion of *GS-quasigroup* was introduced in [5]. It is IM-quasigroup (Q, \cdot) in which the identity $a(ab \cdot c) \cdot c = b$ is satisfied for every $a, b, c \in Q$.

The notion of *quadratical quasigroup* was introduced in [8]. It is IM-quasigroup (Q, \cdot) in which the identity $ab \cdot a = ca \cdot bc$ is satisfied for every $a, b, c \in Q$.

Hexagonal quasigroup is IM-quasigroup (Q, \cdot) in which the additional identity of *semisymmetry*, $ab \cdot a = b$, is valid for every $a, b \in Q$. It was introduced and studied in [7].

Pentagonal quasigroup is IM-quasigroup (Q, \cdot) in which the additional identity of *pentagonality*,

$(ab \cdot a)b \cdot a = b$, is valid for every $a, b \in Q$. It was introduced and studied in [3].

Basic example which motivates the study of IM-quasigroups is $C(q) = (\mathbf{C}, *)$, where $*$ is a binary operation on \mathbf{C} defined by

$$a * b = (1 - q)a + qb,$$

$q \in \mathbf{C} \setminus \{0, 1\}$. This can be rewritten as

$$\frac{a * b - a}{b - a} = \frac{q - 0}{1 - 0}.$$

If we regard complex numbers as the points of the Euclidean plane, this identity means that the points $a, b, a * b$ are the vertices of a triangle directly similar to the triangle with the vertices $0, 1, q$, i.e. the vertices of a triangle of a given form. In hexagonal quasigroups $a * b$ is the third vertex of the equilateral triangle determined by vertices a and b . In pentagonal quasigroups $a * b$ is the third vertex of the regular pentagon determined by vertices a and b . In quadratical quasigroups $a * b$ is the center of the square determined

by a and b . In GS-quasigroups the point $a * b$ actually lies on the line determined by a and b and divides the segment determined by a and b in the golden section ratio.

If we plug equality $a * b = (1 - q)a + qb$ into identities that define certain subclasses of IM-quasigroups ($*$ is now substituted by \cdot), we get equations with variable q . For hexagonal quasigroups we get $q^2 - q + 1 = 0$, whose one solution is $q = \frac{1+i\sqrt{3}}{2}$, exactly the third vertex of the equilateral triangle determined by complex numbers 0 and 1. For pentagonal quasigroups we get the equation $q^4 - 3q^3 + 4q^2 - 2q + 1 = 0$, whose one solution is $q = \frac{1}{4}(3 + \sqrt{5} + i\sqrt{2(5 + \sqrt{5})})$, exactly the third vertex of the regular pentagon determined by 0 and 1. For quadratical quasigroups we get $q^2 - q + \frac{1}{2} = 0$ with a solution $q = \frac{1+i}{2}$, exactly the center of the square determined by 0 and 1. For GS-quasigroups we get $q^2 - q - 1 = 0$ with a solution $q = \frac{1+\sqrt{5}}{2}$, the point that divides the segment determined by 0 and 1 in the golden section ratio.

Actually, more general statements are valid in each of these subclasses of IM-quasigroups, i.e. Toyoda-like representation theorem for that subclass (see [2]). Since all examples in the following sections refer only to hexagonal and pentagonal quasigroups, we state here only theorems for these two subclasses. Similar theorems for quadratical and GS-quasigroups can be found in [8] and [5]. Next two theorems were proved in [7] and [3]. We state them in the form suitable for our considerations.

Theorem 1.1. *There is a hexagonal quasigroup (Q, \cdot) if and only if there is an abelian group $(Q, +)$ with an automorphism φ which satisfies*

$$\varphi^2 - \varphi + 1 = 0. \quad (1)$$

Theorem 1.2. *There is a pentagonal quasigroup (Q, \cdot) if and only if there is an abelian group $(Q, +)$ with an automorphism φ which satisfies*

$$\varphi^4 - 3\varphi^3 + 4\varphi^2 - 2\varphi + 1 = 0. \quad (2)$$

In both of these subclasses operations \cdot and $+$ are related by the equality

$$a \cdot b = a + \varphi(b - a) = (1 - \varphi)(a) + \varphi(b), \quad (3)$$

for $a, b \in Q$.

2. MIDPOINT OF A SEGMENT IN IM-QUASIGROUPS

Elements of a quasigroups are called *points*. A pair of points a and b is called a *segment* and is denoted by $\{a, b\}$. Cyclic n -tuple of points a_1, a_2, \dots, a_n is called *n -gon* and is denoted by (a_1, a_2, \dots, a_n) .

Given four elements a, b, c, d of a medial quasigroup (Q, \cdot) the concept of *parallelogram* was defined in [4]. However, it was necessary to define many nontrivial geometrical concepts in a medial quasigroup in order to define it properly. If we observe medial quasigroups with the additional identity of idempotency, IM-quasigroups, the definition of parallelogram becomes much more elegant (see [6]). In many subclasses of IM-quasigroups this definition can become even more elegant due to an additional identity which determines that subclass. Among such subclasses are four subclasses mentioned in the previous section. Since in this section we only consider examples in hexagonal quasigroups, we only give definition of parallelogram in hexagonal quasigroups (see [9]), but similar definitions can be done in other mentioned subclasses.

Let (Q, \cdot) be a hexagonal quasigroup. We say that a, b, c, d form a *parallelogram* and denote it by $Par(a, b, c, d)$ if $d = bc \cdot ab$ (Figure 1).

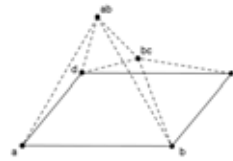


Figure 1: Parallelogram in hexagonal quasigroup

Given the definition of parallelogram the next definition comes naturally. Let (Q, \cdot) be a hexag-

onal quasigroup and $a, b \in Q$. We say that $m \in Q$ is *midpoint* of the segment $\{a, b\}$ and denote it by $M(a, m, b)$ if $\text{Par}(a, m, b, m)$. Using this along with the definition of parallelogram we get that m is the midpoint of the segment $\{a, b\}$ if and only if $b = am \cdot ma$ holds.

In the hexagonal quasigroup $C(q) = (\mathbf{C}, *)$, where q is a solution of the equation $q^2 - q + 1 = 0$, the midpoint m of any segment $\{a, b\}$, $a, b \in Q$ exists, it is unique and we have $m = \frac{a+b}{2}$. We get that easily using $q^2 - q + 1 = 0$:

$$\begin{aligned}
 b &= am \cdot ma \\
 &= (a + q(m - a)) \cdot (m + q(a - m)) \\
 &= (a + q(m - a)) + q((m + q(a - m)) - (a + q(m - a))) \\
 &= a + q(m - a) + q(m - a + 2q(a - m)) \\
 &= a - 2q(a - m) + 2q^2(a - m) \\
 &= a - 2(a - m) = 2m - a.
 \end{aligned}$$

However, that is not the case in the hexagonal quasigroup (Q_4, \cdot) constructed in [1].

Table 1: Hexagonal quasigroup Q_4

\cdot	0	1	2	3
0	0	2	3	1
1	3	1	0	2
2	1	3	2	0
3	2	0	1	3

Let a and b be distinct elements of Q_4 . Let $ab = c$ and $ba = d$. Such elements are for example $a = 0, b = 1, c = 2, d = 3$. We have

$$\begin{aligned}
 aa \cdot aa &= a \cdot a = a \\
 ab \cdot ba &= c \cdot d = a \\
 ac \cdot ca &= d \cdot b = a \\
 ad \cdot da &= b \cdot c = a
 \end{aligned}$$

$$\begin{aligned}
 aa \cdot aa &= a \cdot a = a \\
 ba \cdot ab &= d \cdot c = b \\
 ca \cdot ac &= b \cdot d = c \\
 da \cdot ad &= c \cdot b = d.
 \end{aligned}$$

We notice that for $a \neq b$ there does not exist $x \in Q_4$ such that $M(a, x, b)$, while for every $x \in Q_4$ $M(a, x, a)$ holds. Thus the segment $\{a, a\}$ in Q_4 has four distinct midpoints. The reason behind that lies in the abelian group associated to the quasigroup (Q_4, \cdot) by Theorem 1.1, which is $\mathbf{Z}_2 \times \mathbf{Z}_2$. Using (3) and (1) the identity $b = am \cdot ma$ can easily be transformed into the identity $2m = a + b$, just like in the case $C(q)$. Since there do not exist two different elements of $\mathbf{Z}_2 \times \mathbf{Z}_2$ that add up to the identity element, for $a \neq b$ there does not exist $x \in Q_4$ such that $M(a, x, b)$. Since every element of the group $\mathbf{Z}_2 \times \mathbf{Z}_2$ is of order 2, for every $x \in Q_4$ we have $M(a, x, a)$.

In [1] another example of the hexagonal quasigroup was constructed, so called (Q_3, \cdot) . It is easily checked by direct computation that every segment in Q_3 has the unique midpoint.

Table 2: Hexagonal quasigroup Q_3

\cdot	a	b	c
a	a	c	b
b	c	b	a
c	b	a	c

In the same article the next theorem was proved.

Theorem 2.1. *Let $(Q_1, \cdot_1), (Q_2, \cdot_2), \dots, (Q_n, \cdot_n)$ be hexagonal quasigroups, and let \circ be the oper-*

ation defined on $Q = Q_1 \times Q_2 \times \dots \times Q_n$ by:

$$(x_1, x_2, \dots, x_n) \circ (y_1, y_2, \dots, y_n) = (x_1 \cdot_1 y_1, x_2 \cdot_2 y_2, \dots, x_n \cdot_n y_n).$$

Then (Q, \circ) is a hexagonal quasigroup.

Using the examples Q_3 and Q_4 and the previous theorem the quasigroup $(Q_3 \times Q_4, \cdot)$ can be constructed.

Table 3: Hexagonal quasigroup $Q_3 \times Q_4$

\cdot	a_0	a_1	a_2	a_3	b_0	b_1	b_2	b_3	c_0	c_1	c_2	c_3
a_0	a_0	a_2	a_3	a_1	c_0	c_2	c_3	c_1	b_0	b_2	b_3	b_1
a_1	a_3	a_1	a_0	a_2	c_3	c_1	c_0	c_2	b_3	b_1	b_0	b_2
a_2	a_1	a_3	a_2	a_0	c_1	c_3	c_2	c_0	b_1	b_3	b_2	b_0
a_3	a_2	a_0	a_1	a_3	c_2	c_0	c_1	c_3	b_2	b_0	b_1	b_3
b_0	c_0	c_2	c_3	c_1	b_0	b_2	b_3	b_1	a_0	a_2	a_3	a_1
b_1	c_3	c_1	c_0	c_2	b_3	b_1	b_0	b_2	a_3	a_1	a_0	a_2
b_2	c_1	c_3	c_2	c_0	b_1	b_3	b_2	b_0	a_1	a_3	a_2	a_0
b_3	c_2	c_0	c_1	c_3	b_2	b_0	b_1	b_3	a_2	a_0	a_1	a_3
c_0	b_0	b_2	b_3	b_1	a_0	a_2	a_3	a_1	c_0	c_2	c_3	c_1
c_1	b_3	b_1	b_0	b_2	a_3	a_1	a_0	a_2	c_3	c_1	c_0	c_2
c_2	b_1	b_3	b_2	b_0	a_1	a_3	a_2	a_0	c_1	c_3	c_2	c_0
c_3	b_2	b_0	b_1	b_3	a_2	a_0	a_1	a_3	c_2	c_0	c_1	c_3

If we now take the point $a_0 \in Q_3 \times Q_4$, let us calculate the other endpoint of the segment with the first endpoint a_0 , given all possible midpoints.

$$\begin{aligned} a_0 a_0 \cdot a_0 a_0 &= a_0 a_0 = a_0 \\ a_0 a_1 \cdot a_1 a_0 &= a_2 a_3 = a_0 \\ a_0 a_2 \cdot a_2 a_0 &= a_3 a_1 = a_0 \\ a_0 a_3 \cdot a_3 a_0 &= a_1 a_2 = a_0 \end{aligned}$$

$$\begin{aligned} a_0 b_0 \cdot b_0 a_0 &= c_0 c_0 = c_0 \\ a_0 b_1 \cdot b_1 a_0 &= c_2 c_3 = c_0 \\ a_0 b_2 \cdot b_2 a_0 &= c_3 c_1 = c_0 \\ a_0 b_3 \cdot b_3 a_0 &= c_1 c_2 = c_0 \end{aligned}$$

$$\begin{aligned} a_0 c_0 \cdot c_0 a_0 &= b_0 b_0 = b_0 \\ a_0 c_1 \cdot c_1 a_0 &= b_2 b_3 = b_0 \\ a_0 c_2 \cdot c_2 a_0 &= b_3 b_1 = b_0 \\ a_0 c_3 \cdot c_3 a_0 &= b_1 b_2 = b_0 \end{aligned}$$

Therefore, each of the segments $\{a_0, a_0\}$, $\{a_0, b_0\}$, $\{a_0, c_0\}$ has four midpoints, while each of the segments $\{a_0, a_i\}$, $\{a_0, b_i\}$, $\{a_0, c_i\}$ for $i = 1, 2, 3$ does not have a midpoint. The first statement is the direct consequence of four possible midpoints of the segment $\{i, i\}$, $i = 0, 1, 2, 3$ in Q_4 and the uniqueness of the midpoint in Q_3 . The second statement follows directly from nonexistence of the midpoint for each segment $\{i, j\}$, $i \neq j$, in Q_4 .

3. REGULAR TRIANGLE AND REGULAR HEXAGON IN HEXAGONAL QUASIGROUPS

The next concept was defined in [9]. We say that points a, b, c of the hexagonal quasigroup (Q, \cdot) are vertices of a *regular triangle* if $ab = c$ and we denote it by $Tr(a, b, c)$ (Figure 2). It is clear that the statements $Tr(a, b, c)$, $Tr(b, c, a)$ and $Tr(c, a, b)$ are mutually equivalent and that a regular triangle is uniquely determined by any two of its vertices.

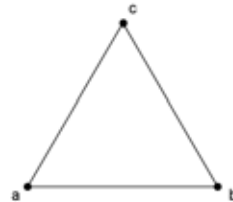


Figure 2: Regular triangle (a, b, c)

We say that points a, b, c, d, e, f of the hexagonal quasigroup (Q, \cdot) form a *regular hexagon* with the *center* $o \in Q$ if $Tr(o, a, b)$, $Tr(o, b, c)$, $Tr(o, c, d)$, $Tr(o, d, e)$, $Tr(o, e, f)$ and $Tr(o, f, a)$ hold. We denote it by $Hex_o(a, b, c, d, e, f)$ or just $Hex(a, b, c, d, e, f)$. It is shown in Figure 3.

We immediately see that regular hexagon (a, b, c, d, e, f) in a hexagonal quasigroup (Q, \cdot) is uniquely determined by its two adjacent vertices, for example a and b . Given $a, b \in Q$ we get

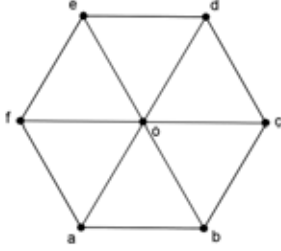


Figure 3: Regular hexagon (a, b, c, d, e, f) with the center o

$o = ab, c = ob, d = oc, e = od, f = oe$. Therefore, we can conclude that in a hexagonal quasigroup center of a given regular hexagon always exists and is unique, because it can be expressed in terms of two adjacent vertices determining that hexagon using operation in the quasigroup.

However, the situation is different with centers of regular triangles. Definition of the center is motivated by Figure 4.

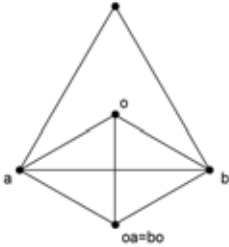


Figure 4: Center of regular triangle

We say that $o \in Q$ is the *center* of regular triangle (a, b, c) in the hexagonal quasigroup (Q, \cdot) if $oa = bo$ holds.

In the quasigroup $Q_3 = \{a, b, c\}$ the statement $Tr(a, b, c)$ holds. If we try all possible candidates for the center o of (a, b, c) we get

$$aa = a \neq ba, ba = c \neq bb, ca = b \neq bc.$$

Therefore, we can conclude that regular triangle (a, b, c) in Q_3 does not have a center. The reason

behind that lies in the use of Theorem 1.1, precisely in the use of equations (3) and (1) on the identity $oa = bo$:

$$\begin{aligned} (1 - \varphi)(o) + \varphi(a) &= (1 - \varphi)(b) + \varphi(o) \\ (1 - 2\varphi)(o) &= (1 - \varphi)(b) - \varphi(a). \end{aligned}$$

The quasigroup Q_3 is associated to the abelian group \mathbf{Z}_3 with the automorphism $\varphi(x) = 2x$ which clearly satisfies (1). Since the function $1 - 2\varphi = 0$ is not invertible, there does not exist $o \in Q_3$ satisfying

$$(1 - 2\varphi)(o) = (1 - \varphi)(b) - \varphi(a).$$

We can immediately notice that if the center o of a regular triangle (a, b, c) in a hexagonal quasigroup (Q, \cdot) exists, it is unique, since

$$o = (1 - 2\varphi)^{-1}((1 - \varphi)(b) - \varphi(a))$$

holds.

4. REGULAR PENTAGON AND REGULAR DECAGON IN PENTAGONAL QUASIGROUPS

We say that points a, b, c, d, e of the pentagonal quasigroup (Q, \cdot) form a *regular pentagon* if $ab = c, bc = d$ and $cd = e$ and we denote it by $RP(a, b, c, d, e)$ (Figure 5). It can easily be shown that $RP(a, b, c, d, e)$ implies $de = a$ and $ea = b$ and that relations $RP(a, b, c, d, e)$, $RP(b, c, d, e, a)$, $RP(c, d, e, a, b)$, $RP(d, e, a, b, c)$ and $RP(e, a, b, c, d)$ are mutually equivalent. It is also clear from the definition that a regular pentagon is uniquely determined by the pair (a, b) of its vertices.

We say that points $a, b, c, d, e, f, g, h, i, j$ of the pentagonal quasigroup (Q, \cdot) form a *regular decagon* with the center $o \in Q$ if

$$\begin{aligned} ba \cdot b &= cb \cdot c = dc \cdot d = ed \cdot e = fe \cdot f = \\ gf \cdot g &= hg \cdot h = ih \cdot i = ji \cdot j = aj \cdot a = o \end{aligned}$$

We denote it by $RD_o(a, b, c, d, e, f, g, h, i, j)$ or just $RD(a, b, c, d, e, f, g, h, i, j)$. It is shown in Figure 6.

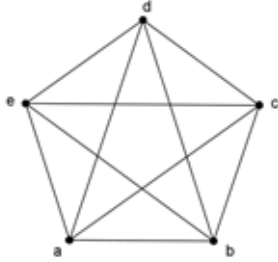


Figure 5: Regular pentagon (a, b, c, d, e)

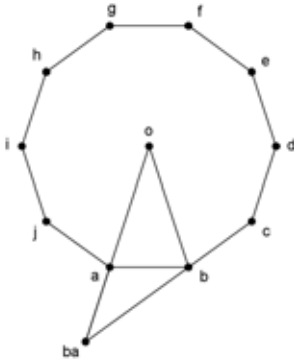


Figure 6: Regular decagon $(a, b, c, d, e, f, g, h, i, j)$ with the center o

We easily see that regular decagon in a pentagonal quasigroup (Q, \cdot) is uniquely determined by its adjacent vertices a and b . Given $a, b \in Q$ we get $o = ba \cdot b$, $j = ba \cdot a$, $i = bo$, $h = ao$, $g = jo$, $f = io$, $e = ho$, $d = go$, $c = fo$. Therefore, we can conclude that in a pentagonal quasigroup center of a given regular pentagon always exists and is unique, because it can be expressed in terms of two adjacent vertices determining that pentagon using operation in the quasigroup.

The situation is different for regular pentagons. Definition of the center is motivated by Figure 7.

We say that $o \in Q$ is the *center* of regular pentagon (a, b, c, d, e) in the pentagonal quasigroup (Q, \cdot) if $oa \cdot b = o$ holds.

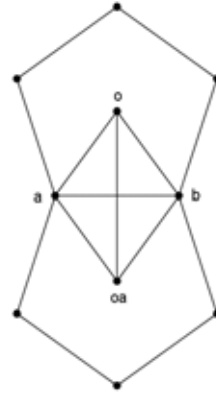


Figure 7: Center of regular pentagon

Let us now consider the pentagonal quasigroup (Q_5, \cdot) .

Table 4: Pentagonal quasigroup Q_5

\cdot	0	1	2	3	4
0	0	2	4	1	3
1	4	1	3	0	2
2	3	0	2	4	1
3	2	4	1	3	0
4	1	3	0	2	4

In (Q_5, \cdot) the statement $RP(0, 1, 2, 3, 4)$ holds. The pentagon $(0, 1, 2, 3, 4)$ is determined by its vertices 0 and 1. Now we want to try all possible candidates for the center o of $(0, 1, 2, 3, 4)$:

$$00 \cdot 1 = 2 \neq 0, 10 \cdot 1 = 3 \neq 1, 20 \cdot 1 = 4 \neq 2, \\ 30 \cdot 1 = 0 \neq 3, 40 \cdot 1 = 1 \neq 4.$$

Hence, we conclude that the regular pentagon $(0, 1, 2, 3, 4)$ in Q_5 does not have a center. Algebraic justification for that is the use of Theorem 1.2. If we apply equations (3) and (2) on the identity $oa \cdot b = o$, we get successively:

$$(1 - \varphi)(oa) + \varphi(b) = o$$

$$\begin{aligned}
(\mathbf{1} - \varphi)^2(o) + (\mathbf{1} - \varphi)\varphi(a) + \varphi(b) &= o \\
o - 2\varphi(o) + \varphi^2(o) - o &= -\varphi(\mathbf{1} - \varphi)(a) - \varphi(b) \\
\varphi(2 \cdot \mathbf{1} - \varphi)(o) &= \varphi(\mathbf{1} - \varphi)(a) + \varphi(b).
\end{aligned}$$

Since φ is an isomorphism it is equivalent to

$$(2 \cdot \mathbf{1} - \varphi)(o) = (\mathbf{1} - \varphi)(a) + b.$$

The quasigroup Q_5 is constructed from the abelian group \mathbf{Z}_5 using the automorphism $\varphi(x) = 2x$ which clearly satisfies (2). Function $2 \cdot \mathbf{1} - \varphi = 0$ is not invertible, so there does not exist $o \in Q_5$ that satisfies

$$(2 \cdot \mathbf{1} - \varphi)(o) = (\mathbf{1} - \varphi)(a) + b.$$

From the last equation one can immediately notice that if the center o of a regular pentagon (a, b, c, d, e) exists, it is unique, because

$$o = (2 \cdot \mathbf{1} - \varphi)^{-1}((\mathbf{1} - \varphi)(a) + b)$$

holds.

5. CONCLUSIONS

In this paper the concepts of IM-quasigroup and some of its subclasses are defined. The special emphasis is put on hexagonal and pentagonal quasigroups. In these subclasses some geometrical concepts are defined: midpoint of a segment, regular triangle and regular hexagon in hexagonal quasigroups and regular pentagon and regular decagon in pentagonal quasigroups. The existence and uniqueness of the midpoint of a segment and of the centers of regular triangles and regular hexagons is studied in hexagonal quasigroups and the existence and uniqueness of the centers of regular pentagons and regular decagons is studied in pentagonal quasigroups.

It turns out that a segment in a hexagonal quasigroup does not necessarily have to have a midpoint, while some segments can have multiple midpoints. It also turns out that the centers of regular triangles and pentagons do not have to exist, while the centers of regular hexagons and decagons always exist and are unique in

these subclasses of IM-quasigroups. Some examples of finite quasigroups are given to illustrate these results. The algebraic justification of the mentioned results using characterization of these subclasses via abelian groups which possess a certain type of automorphisms is also given. For hexagonal quasigroups these are the automorphisms φ which satisfy functional equation $\varphi^2 - \varphi + \mathbf{1} = 0$ and for pentagonal quasigroups the automorphisms φ which satisfy $\varphi^4 - 3\varphi^3 + 4\varphi^2 - 2\varphi + \mathbf{1} = 0$.

REFERENCES

- [1] M. Bombardelli. Finite hexagonal quasigroups. *Quasigroups Related Systems*, 14: 157–162, 2006.
- [2] K. Toyoda. On axioms of linear functions. *Proc. Imp. Acad. Tokyo*, 17: 221–227, 1941.
- [3] S. Vidak. Pentagonal quasigroups. *Quasigroups Related Systems*, 22: 147–158, 2014.
- [4] V. Volenec. Geometry of medial quasigroups. *Rad JAZU [421]*, 5: 79–91, 1986.
- [5] V. Volenec. GS-quasigroups. *Čas. Pěst. Mat.*, 115: 307–318, 1990.
- [6] V. Volenec. Geometry of IM-quasigroups. *Rad HAZU [456]*, 10: 139–146, 1991.
- [7] V. Volenec. Hexagonal quasigroups. *Archivum Mathematicum (Brno)*, 27a: 113–122, 1991.
- [8] V. Volenec. Quadratical quasigroups. *Note di Matematica*, 13: 107–115, 1993.
- [9] V. Volenec. Regular triangles in hexagonal quasigroups. *Rad HAZU [467]*, 11: 85–93, 1994.

ABOUT THE AUTHORS

Stipe Vidak obtained his undergraduate degree in mathematics in 2004 and his PhD in mathematics in 2012 from University of Zagreb. He

works as a research and teaching assistant at Department of Mathematics, University of Zagreb, Croatia. His main fields of interest are nonassociative algebraic structures and their applications in geometry.

EXPERIMENTS WITH A FOLDING MULTI-AGENT SYSTEM IN THE DESIGN OF TRIANGLE MESH STRUCTURES

Benjamin FELBRICH¹, Daniel LORDICK², Jörg Rainer NOENNIG¹ and
Sebastian WIESENHUETTER¹

¹TU Dresden, Germany, Dept. of Architecture, Juniorprofessorship of Knowledge Architecture

²TU Dresden, Dept. of Mathematics and Science, Institute of Geometry

ABSTRACT: In the presented work we attempt to find a way to utilize the collaborative intelligence of bio-inspired multi-agent systems for an architectural application. To do so we establish, geometrically describe, optimize and program a physical object. Our setup is a foldable structure of interconnected triangles where simple robotic agents control the angles between adjacent facets. This structure has the capacity to reshape itself in a self-directed manner. The preliminary objective was to let the structure approximate a certain target shape. Beyond this we investigate more general concepts for the application of the system. Apart from the geometrical foundation and the computer simulation of the folding behavior, the actual prototyping with physical models plays a significant role.

Keywords: prototyping, rigid folding, triangle mesh, free form discretization

1. INTRODUCTION

In nature, systems consisting of a large number of simple individuals achieve stunning goals by interacting in both physical and cognitive ways. Both ways, from collaboratively lifting heavy objects to the optimization of food paths, have in common that the abilities, which emerge from the group, go far beyond the abilities of an individual actor (“agent”).

We tried to find a mechanism resp. construction kit which utilizes the logic of emergence for a defined architectural purpose – the erection of free form structures. Composed from a multitude of simple robotic agents, the structure is supposed to reshape itself and thereby approximately regenerate an arbitrarily chosen free form surface. Such a device would be able to directly construct a form imagined by the architect, or to generate a form which is optimal under certain conditions. Furthermore it could be used as a design tool, which gives a direct physical feedback to form alterations. In

larger scale applications such robot driven structures could serve as a flexible concrete formwork: after assembling a particular configuration of triangular agents, a set of free form shapes within a certain design space can be erected.

Upon close investigation, a folding mechanism appeared promising, in which each folding edge is controlled by a single agent and all interconnected parts form a large dynamic system, a discrete surface. The distribution and the alignment of the agents were the first important issue to be closely investigated. This topic belongs to the realm of free form discretization and optimization, since the mechanism’s underlying folding pattern is derived from the triangulation of the target surface to regenerate. The other important part of the project was to elaborate an agent design that is capable of lifting weight by rotating adjacent frames. It was developed with the help of physical prototypes, which were tested, evaluated and enhanced step by step.

In the following, both parts of the work will

be shown. In the end we will present an experimental setup in which the device has to reach a predefined target surface. Prior to this, some preconditions need to be clarified.

2. ASSUMPTIONS

Our approach necessitates some characteristics which correlate with cost-relevant requirements of real-world applications like standardization. Beside this positive aspect the system nevertheless should offer a rather broad range of achievable shapes. In detail we make the following assumptions to assure feasibility:

- **Simplicity** - the system and its parts have to be as simple as possible
- **Similarity** - each robotic agent has to resemble its neighbor
- **Interdependence** - to reach the common goal, each agent has to actively participate
- **Smoothness** - the triangulated structure has to be as smooth and as close to the target surface as possible
- **Flexibility** - the system has to have the ability to approximate any free form surface

To achieve the latter the structure is made of a set of triangular parts with a limited variety. The configuration of the triangles can be changed manually to access different design spaces.

3. DEVELOPING PROCESS

The consecutive steps in the process of developing the structure and the algorithms to control the structure, were as follows: starting from the design of a free form surface, the mechanism was constructed to approximate this designed target surface. Then the mechanism was folded into an arbitrary state. From this state it had to restore the initial constitution by itself. After this goal was reached, the target surface can be changed. Now the construction kit can be reconfigured and thus be adapted to the new

conditions – without changing the agents' design but only their topological alignment.

4. TRIANGULATION

The abstraction of the described folding setup is represented by a triangular polygon mesh, in which the folding agents are located on the inner mesh edges.

Since we intend to establish a construction kit, which is affordable and easy to produce, we have to maximise the degree of repetition of components the system is made of. A conventional free form triangulation, which in general produces nothing but unique edge lengths, would have forced us to individually draw and produce every piece of the frame. Though this process can be handled by means of parametric modelling and tools like computer aided manufacturing (CAM), the re-configurability would be lost.

A more feasible approach was found in limiting the amount of utilized edge lengths in the triangulation to a relatively low number. Thus a limited set of combinable triangles emerges that can be easily reproduced.

In order to approximate any free form surface we had to investigate the relation between smoothness and standardization and derive an algorithm to triangulate free form surfaces using a limited amount of edge lengths. It is inherent to this problem that we are not able to achieve a triangulation result in which every vertex of the mesh is placed on the target surface S or in which every edge precisely equals an item of the limited set of edge lengths yet certain deviations may be accepted in this regard.

4.1 Discretization – Standardization versus Smoothness

It is save to say that complexity C increases with the number of edge lengths utilized in the triangulation. C needs to be minimized. Smoothness on the other hand has a close rela-

tion to the discrete Gaussian curvature. In [1] Alexander Bobenko defined the discrete Gaussian curvature K in a node p of a discrete surface S as:

$$K(p) := 2\pi - \sum_i \alpha_i$$

Here α_i are the angles between nearby edges adjacent to p .

In the simplest setup we can assume all edge lengths to be equal and thus form only equilateral triangles. This setup provides the highest degree of repetition, i.e. C cannot be lowered any further. In this case we have distinct steps of discrete Gaussian curvature depending on the number of edges adjacent to p . A node p with n adjacent edges will have the curvature:

$$K(p) = 2\pi - n \cdot \frac{\pi}{3}$$

This results in a very limited variety of curvature that can be obtained with one edge length available. We consider the number of alternatives for the discrete Gaussian curvature in each node to be the graduator for the obtainable smoothness in the triangulation. In order to increase this smoothness but still keep a high degree of repetition in the utilized pieces, we allowed the triangulation algorithm to use a limited set of edge lengths larger than one. Doing so, the number of alternatives of discrete Gaussian curvature per node increases exponentially with the number of available edge lengths.

To allow the system to move more freely, we perforated the mesh by removing certain edges. Physical mockups showed that a favorable folding behavior can be achieved if every node of the mesh is either a border node itself, or if it is neighbor to at least one border node. As abbreviation, the described mesh will be called a “Perforated Limited Triangular Mesh” (PLT-Mesh) with a rank $|L|$, in which $|L|$ describes the number of different edge lengths used. Figure 1 shows an example of a

PLT-Mesh of the rank 4.

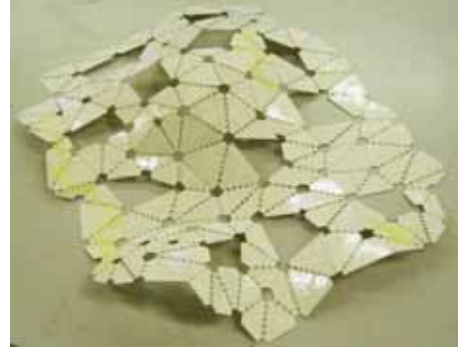


Figure 1: PLT-Mesh of 4th rank

4.2 Triangulation algorithm

To harness the theoretical potential of the limited mesh topology in smoothness and standardization, it was essential to find a way to properly discretize the given free form surface S only using l_i . The algorithm established for that purpose is as follows:

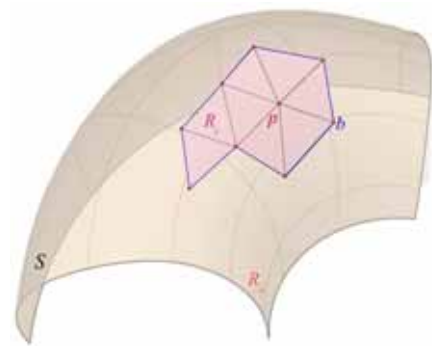


Figure 2: Definitions

Definitions:

S Surface to triangulate

- L a finite set of previously chosen values for edge lengths with $|L|$ being its cardinality
- l_i the elements of L
- p nodes of the polygon mesh to be compiled
- b boundary of the polygon mesh, the closed polyline consisting of all the mesh edges with only one adjacent mesh face
- b_p normal projection of b on S
- R_i the region of S that is inside b_p
- R_o the region of S that is outside b_p

Basic algorithmic procedure:

1. Set one initial triangle arbitrarily with the condition that all of its three vertices are placed on S and only l_i are used as edge lengths.
2. Choose one of the triangle's edges and find two new lines (again only using l_i) to form a new neighbor triangle on S . (by creating spheres with the radii l_i around the respective nodes. One of the two intersection points x_1 and x_2 of these spheres and S is the new node.

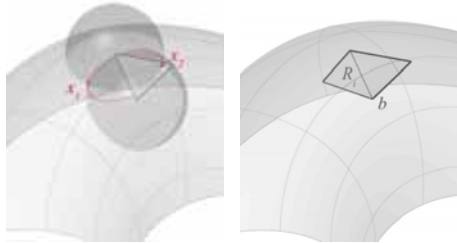


Figure 3: Steps two and three of the triangulation algorithm

3. Include the new triangle and update the polygon mesh and b resp. b_p
4. Examine b and pick a single line segment or two adjoining segments of b to operate on. Then pick and execute one of the following options*:
- 4.a) Repeat step 2 with a single segment of b (which adds one new triangle).

- 4.b) If two adjoining line segments of b are chosen, fill the narrow gap they span over R_o with two new triangles. This can be geometrically solved by creating $|L|$ spheres with radii l_i around each of the three nodes of the chosen polyline segment. The intersection points of these spheres form a point cloud in which every point has the property of correct distances from the adjacent nodes. From this cloud of points, pick the one with the shortest distance to its normal projection on S and with its normal projection on S being outside R_i . This point represents the new node to add two triangles.

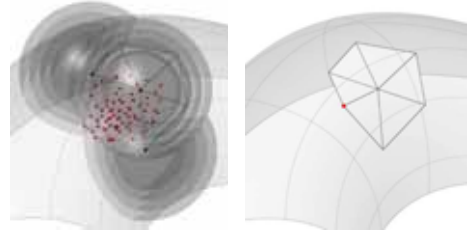


Figure 4: Step 4.b) of the triangulation algorithm

- 4.c) If the angle between two neighbor segments over R_o falls below a certain threshold value, directly connect the segments' two outer nodes with a new line.
5. If S is filled with triangles, stop. Otherwise go back to 3.

* The conditions under which the operation mode in 4. is chosen represents a complex issue of optimizing the algorithm's performance. The priority is to find the narrowest angle over R_o between adjoining segments of b and close them using mode 4.b) preferably or 4.c) in critical situation. These conditions are inspired by and similar to the behavior that wasps show when choosing the most enclosed comb cell to fill when building their nests [2]; it is crucial for avoiding self intersections in b_p and make the triangles grow evenly to the outside over S .

The described algorithm was coded and executed using Grasshopper in Rhino 5. It produces inevitable inaccuracies since it partly draws new nodes near S but not necessarily places them on S . Thus the resulting mesh is a rather bumpy approximation of S . This geometrically inherent error can be partly or completely assigned to the accuracy of the edge lengths by replacing them with their normal projection on S (very inaccurate) or pulling them towards S using a solver implemented in the particle spring engine Kangaroo (more accurate). Thus it is possible to gradually increase smoothness if a certain tolerance in edge length deviation is allowed.

With this algorithm we found a way to optimally place the individual components of the multi-agent system on a free form.

5. PROTOTYPING

The following paragraphs describe, how the agents were designed and how the prototypical test structure was executed. Looking for different ways to empower the agents, a mechanism with a spindle drive appeared most suitable, in which a rotating shaft pushes or pulls pieces of the frame apart or towards each other. This option was chosen due to its light weight, its low cost and the favorable power transmission. Figure 5 shows the basic idea; figure 6 the section of the frames indicating the most important parameters.

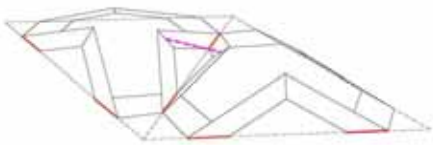


Figure 5: Basic idea; red: hinges, magenta: shaft to push / pull the frame

After determining the agent's basic design in

drawings and simple mockups, three consecutive prototypes (so-called "Kingfisher") were elaborated. To quickly proceed in their evolution, small groups of agents (up to three) were built, tested, and improved in regards to their strength and robustness. This method of operation represented an important shift in our approach from a merely theoretical, simulation-based mode to the very practical design-oriented model making. It was inspired by the shifts in innovation emergence in a design-context described in [3].

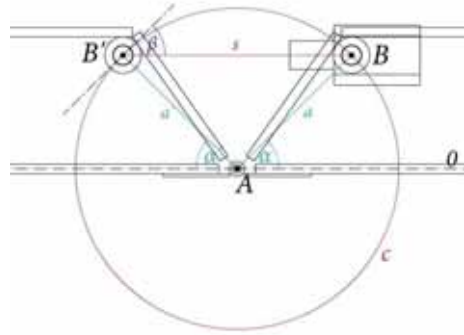


Figure 6: Basic section

5.1 Kingfisher 0.1

The very first prototype (figure 7) consisted of a MDF frame, metallic hinges and was empowered with a 12V Motraxe X-Slot Race 143 model train motor with an idle speed of about 22000 rpm and a torque of 1.3 Nmm (manufacturer information). It was controlled using an Arduino Uno and the Firefly plug-in for Grasshopper in Rhino 5. The major insight from this model was that the utilized MDF was too heavy for the relatively weak motor. Furthermore it revealed the major importance of the shield angle α (see figure 6), since it directly affects the leverage and therefore the power transition from the motor's rotation force into a push / pull force along the threaded shaft

s into a rotation of the frames about their common edge A.



Figure 7: Kingfisher 0.1

5.2 Kingfisher 0.2

In the second prototype shown in figure 8 the shield angle α was optimized to improve the agility by a better power transmission. Further, a lighter foam board material and the stronger motor “Motraxe X-Slot 283 Tuning” (4.0 Nmm torque) were utilized.

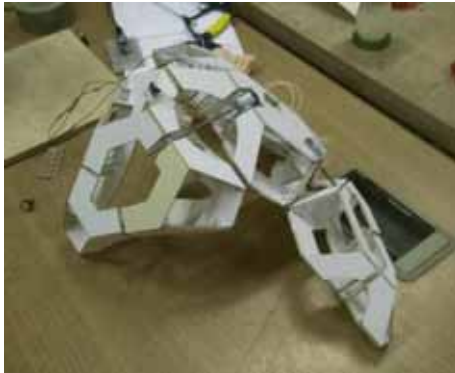


Figure 8: Kingfisher 0.2

The mount of the motors and their counter pieces were improved just like the overall precision of the model making. Like in its predecessor, the motors were controlled using an

Arduino Duemilanove and the interface “Fire-fly” for Grasshopper in Rhino 5. Furthermore a motor control unit based on the L293 chip was used with an external power source to turn the relatively low 5V control voltage of the Arduino board into a powerful 12V operating voltage with a current of up to 500mA per channel to empower up to two motors. In contrary to its predecessor this prototype was able to lift itself off the ground and showed a much higher ability to move in general.

5.3 Kingfisher 1

In the third prototype (figures 9 to 13) major improvements were made: the empowering unit was separated from the triangular frame which interconnects the agents. This was done to ensure its re-configurability (see section “2. Assumptions”). It is furthermore equipped with a gear box with a transition ratio of 1:4 (figure 9), a stronger motor (“Johnson 20543” with 130 Nmm torque) and a potentiometer mounted to the frames giving a feedback of the rotation and angular state in the agent. Its highly stressed parts (e.g. the mount of the motors) are made of stable MDF while the rest is made of lighter foam board.

With a shield angle α of 47.6° the agent is able to rotate its counter part 62.5° in the upper direction until the agent is fully closed (figure 12). It can also rotate its counter part 62.5° in the lower direction until the operating angle between shaft and the rotation circle’s tangent equals 75° . In this fully opened state (figure 13), about 0.96 times the motor’s force is pushing and pulling the opposite pieces apart instead of making them rotate about A. This value was estimated as the upper bound for stress in the frame in this direction, known from the previous models. To prevent the frames from cracks, these values were embedded in the code for the motor not to exceed this area of operation.

In this prototype we used motor control shields based on the L298N chip to receive a powerful operating voltage of 12V and up to 2A current per channel (two channels).



Figure 9: Gear box of Kingfisher 1



Figure 10: Kingfisher 1, single agent

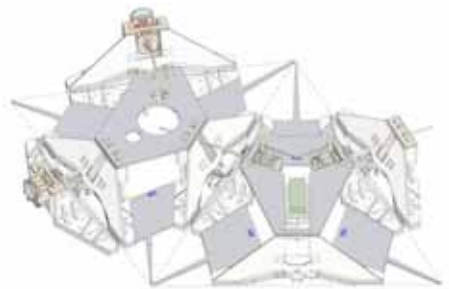


Figure 11: Four agents mounted to connect-

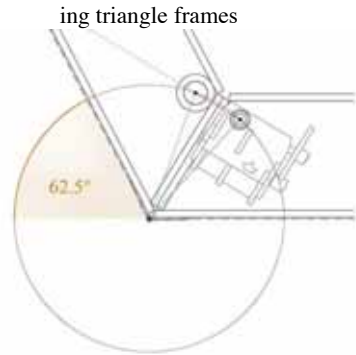


Figure 12: closed state

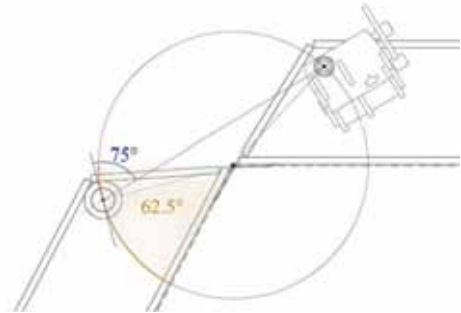


Figure 13: opened state

6. EXPERIMENTAL SETUP

The experimental setup represents the synthesis of a singular agent's design and the logics of the connectivity of a set of agents. It was chosen big enough to provide a sufficiently number of agents – to give a reliable feedback over constraints given by the dimensions. On the other hand it had to be affordable and constructable within only six weeks.

A free form surface was designed in Rhino 3D and triangulated with the algorithm described in section 4.2. Thereupon its appearance could be changed through a simulated folding process using Grasshopper in Rhino 5 and the physics engine Kangaroo. Figures 14 a)

to d) show the initial target surface (a), its discretization into a PLT-Mesh of 3rd rank in b) and another state into which it was transformed through folding in c), whereas d) shows the physical model in this state. There are 22 inner edges of which 21 are equipped with agents. The prototype's dimensions are about 1.10m by 1.40m and a height of about 0.50m



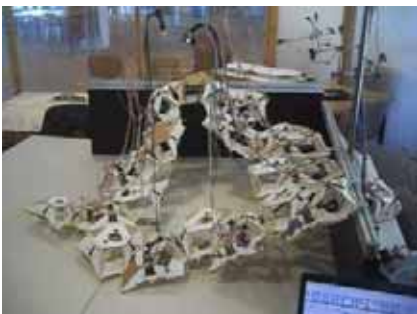
a)



b)



c)



d)

Figure 14: Physical model and test setup

6.1 Electronic Setup

motors: 21 x Johnson 20543, Murata Z5U-5 ceramic capacitor 1 μ F (21 times)
 motor driver: 12 x L298N two channels, up to 2A each
 potentiometer: 21 x Piher PT-15 10 k Ω
 control: 1 x Arduino Duemilenove (for global control), 11x SparkFun Pro Micro 5V 16MHz
 power supply: 2x LPK2-23 400W power supply, 12V 22A each

One laptop is used to send global ON/OFF and timing signals via Firefly to the Arduino Duemilenove. Then these signals are distributed to the smaller agents' boards SparkFun Pro Micro.

6.2 Coding

While the Duemilenove runs with the Firefly Firmata, the SparkFun Pro Micro boards were equipped with a custom code which utilized the feedback of the potentiometers and global time and switch signals.

In order to reduce cost, not every agent was equipped with one control board. Instead, one control board controls two motors. In the code, however, their circuits run completely independent mimicking an independent behavior of two agents.

The goal of the mechanism was set for it to reach its target shape (see figure 14 b) from any starting position (e.g. figure 14 c). To do so every agent has to reach its individual target angle value which was implemented in the code. When every agent has reached this value, the connected mechanism has the state of the target surface.

Two different controls were tested:

A. Straight forward approach:

1. Compare the potentiometers value to your individual target value;
2. If it doesn't equal, rotate the motor towards the target value

3. If you have reached the target value, stop;

Since the interconnected agents naturally block each other with a simple control mode like this, the system would get stuck and even crack. Previous simulations of the folding process predicted these problems. To avoid this, a more sophisticated simulation was run in which the starting position was well defined, the target surface was reached through folding and no deadlocks and cracks appeared. From this simulation the progressive graph of the angle over each inner edge was recorded. Thus, individual control curves for each agent could be established, translated into an array of 10-bit integers and implemented in their individual codes. Following these individual control curves, the agents would precisely repeat the folding simulation in which they do not block each other and finally reach their goal. For this setup a synchronization through a global timing signal was necessary. The simplified commands look like this:

1. Go to the very first value in your control curve (meaning: reset to starting position)
2. If a certain time interval is exceeded, move to the next value in the control curve.
3. Move towards this control curve value.
4. Go to 2.

Here the choice of the time interval represents an issue of fine tuning.

6.3 Test Results

The two different code setups cause very different outcomes. While the straight forward approach makes the mechanism move quickly but uncoordinated and is also more likely to cause cracks (it was more time spent repairing the model than testing it) the second approach led to a more diffident but relatively weak movement.



Figure 15: lower state

In both cases the robot was not able to completely reach its goal due to a lack of strength when lifting heavy parts. Even worse, instead of completely lifting itself into its desired position, it cracked parts of the frame. Figure 15 shows its lower, most unstressed position (resp. figure 14 c) while figure 16 shows its best attempt to reach the target position (resp. figure 14 b) achieved with the first code setup (straight forward). Despite this failure the robot shows a very agile movement within its range of abilities.



Figure 16: upper state

7. CONCLUSION

The initial goal for the device was to reach a

previously defined free form target surface from any initial state and to switch between those states. By way of correct alignment of the pieces in a triangular arrangement it was ensured that – geometrically – the mechanism was able to perform the move. Furthermore the simulation of the desired folding behavior was relatively easy to carry out. In its physical representation, however, the robot failed to reach its goal. This twofold result reveals an advantage of our approach to use both computer aided design and simulation plus physical prototyping for discovering the differences and difficulties that distinguish real world application from its theoretical foundations.

The lack of strength and stability does not only call for stronger frames and motors, but for a more sophisticated and intelligent control mode, which would enable the robot to reach its goal even with limited physical abilities. Such a control, based on agent communication and collaboration would correspond to the natural model of a multi-agent system and be a truly bio-inspired intelligent device. From this point of view, the mechanism presented can act as a test setup for different codes. It provides a physically limited body, comprised of independent entities, that serves as a nucleus for the testing of the performance of collaborative codes.

In extrapolation of the potentials of the mechanism we see a structure that in future can respond to physical interaction with optimization routines and thus can be a sophisticated raw material for design studies. This may push the idea of prototyping within the realm of architecture onto the next level.

REFERENCES

- [1] Alexander I. Bobenko, Yuri B. Suris, “Discrete Differential Geometry: Integrable Structure”, Graduate Studies in Mathematics, Vol. 98, AMS, 2008, page 27-28
- [2] Eric Bonabeau, Marco Dorigo, Guy Theraulaz, “Swarm Intelligence - From

Natural to Artificial Systems”, A volume in the Santa Fe Institute Studies in the sciences of complexity, 1999, Oxford University Press, page 14-23

- [3] Julian Adenauer, Jörg Petruschat: Prototype! - physical, virtual, hybrid, smart - tackling new challenges in design and engineering, form+zweck, 2012, page 12 - 37

ABOUT THE AUTHORS

1. Benjamin Felbrich graduated from the Department of Architecture of TU Dresden. He focuses on the improvements of architectural design through computation and optimization.
2. Daniel Lordick is Professor for Geometric Modelling and Visualization at the Department of Mathematics and Science of TU Dresden, Germany.
3. Jörg Rainer Noennig is Juniorprofessor for Knowledge Architecture at the School of Civil and Environmental Engineering of TU Dresden, Germany.
4. Sebastian Wiesenhütter is an architecture graduate and currently works as a research associate at the chair of Knowledge Architecture at TU Dresden. He focuses on bio-inspired architecture and generative approaches in design.

EXPLORING THE COMPLEXITY. DIGITAL TURN TOWARDS GEOMETRY IN CONTEMPORARY ARCHITECTURE AND URBAN PLANNING

Paweł RUBINOWICZ

West Pomeranian University of Technology Szczecin, Poland

ABSTRACT: The article examines concept of complexity in architecture and urban planning presenting proposal of systematics within the four methods describing complex forms: decomposition, deformation, dispersion and new complexity. The principles of each method are defined on general level. However, within each of the methods, more complex simulations, possible for unequivocal geometrical encoding, can be formulated. The methods can be applied for creation or analysis of both, abstract (2D and 3D composition), and real (specific building or city) forms. The article puts also under investigation application of these methods: for purpose of interpretation and systematics of contemporary architectural creation, development of new tools in design and for better diagnosis of city development principles. In recent years new tools enabling application of advanced digital techniques for exploration of unique formal solutions in architecture and new analytic methods for urban structures have appeared. The geometry becomes more and more universal language important for establishing bases for new digital techniques and interpretation of structure of complex forms in architecture and urban planning.

Keywords: contemporary architecture, urban analysis, fractal geometry, deterministic chaos, 3D virtual city models.

1. INTRODUCTION

Contemporary digital tools create new possibilities for developing buildings and cities. They so broaden aesthetic awareness. Using computers in the architectural and urban planning is already a well-established process. It became popular in 1980s and 90s. Initially, the application of new techniques boiled down to changing from analogue to digital tools. Although the new environment influenced quality and efficiency of work, it did not bring new values to designing, similarly to replacing a typewriter with a computer text editor did not influence the quality of content. Recent years witnessed a major change. New tools and possibilities of using more advanced digital techniques become popular which enabled wider use of computer's capacity. This has been followed by a search for unique formal solutions

in architecture and new methods of analyzing the city space.

Catalysts of the new development in architecture include parametric modeling, NURBS and new prefabrication techniques: possibility of recording a digital form of a facility (CAD) as a sequence of instructions (CAM) recognized by CNC machines, dynamic development of techniques and growing accessibility of 3D print, etc. New technologies enable creating complex forms and reducing building cost. A sign of those changes is growing interest among young architects in programming, creating scripts (e.g. Grasshopper) and consequently more advanced form coding. Although the interest in complexity has been present in architecture since deconstructivism (1980s) and in its primary phase originated from philosophy and literature, digital tools are the driving force of the contemporary development.

However in urban planning, the recent tool is 3D virtual city model. Development in geo-information research, airborne laser scanning techniques (LIDAR), aerial photography analytic techniques, make possible automation and significant acceleration in process of generation virtual city models and new standards of encoding urban structures (CityGML) – cause, that accessibility and accuracy of 3D models is increasing in geometric progression. Sitting at the desk and staring at computer screen, we can easily browse virtual landscapes of most of world's agglomerations, using common free software of Google Earth. The applied tool extends possibilities of simple visualization. 3D city models are more often a medium for application of advanced urban analyses, impossible to execute without computer support.

Both, in architecture and urban planning, the subject of creation or analysis is complexity. However, geometry is universal code necessary for establishing bases of new digital techniques and for interpretation of complex forms. The article presents proposal of classification of this problem in reference to architecture and urban planning.

2. METHODS FOR DEVELOPING COMPLEX FORMS

The chapter presents proposed classification of possible methods for developing complex forms. Their hierarchy was created to interpret contemporary architectural projects. Methods of decomposition, deformation, dispersion and new order are described by the author in his PhD thesis [15]. The description is general enough so it can apply to various real or abstract spatial compositions, including urban structures. A form is understood as a specific shape, and the methods describe the construction process. Definitions of the methods determine pattern and scope of possible geometrical operations (measures). However, each method allows for more detailed simulations that provide for unequivocal, strict geometrical description. Presuppositions of those methods are

presented in fig. 1, below. An example of compositions developed can be seen in fig. 2.

2.1 Decomposition

A set of possible operations within the mechanism of the method includes: break, crush, divide, separate, and fragmentation. The goal is to break the initial form or free compilation of various forms. The main feature of the method is purposeful breaking and far reaching transformation of a structure leading to new composition and aesthetic values. It is based on two essential measures: A) dividing the form and separating specific parts of it in a way which does not result from its construction (form organization logic); B) composing a form using independent parts – not by matching their shapes but by clashing them and allowing mergers. Decomposition can be used to rebuild the initial form (first A, and then B) or to develop a new form (only B). The aim is to highlight the heterogeneity of its construction, based on strong articulation of diversity of its parts and their geometrical independence. Measure A is presented in fig. 1. An example of B is simple overlapping of two regular meshes, which leads to a much more complex form (fig. 3).

2.2 Deformation

The method aims at plasticizing and transforming the structure of a form and at the same time preserving its indivisibility. Operations or terms typical for the method include the following: link, connect, fold, unification, etc. Presuppositions of the method are associated with mathematical topology. Composition measures concentrate on the form, treated as a whole, contained in one specific shape with all constituent parts subordinated to it. The construction process involves two measures: A) transformation of the initial form while preserving links between its parts; B) making the form of its parts – by their fluid merger leading to their mutual unification. The method can be used for rebuilding the initial form (measure A) or building a new form (measure B). The goal

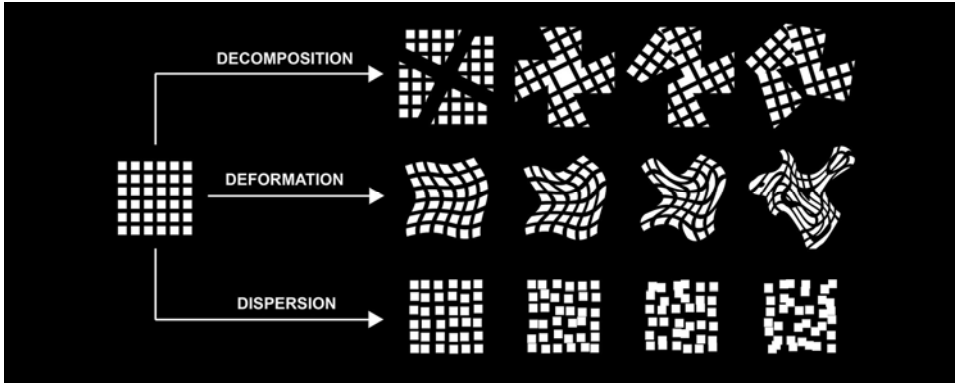


Figure 1: Schemes presenting methods of decomposition, deformation and dispersion



Figure 2: Compilation of schemes presenting creation of complex forms with examples of spatial arrangements. From left: methods of decomposition, deformation and dispersion

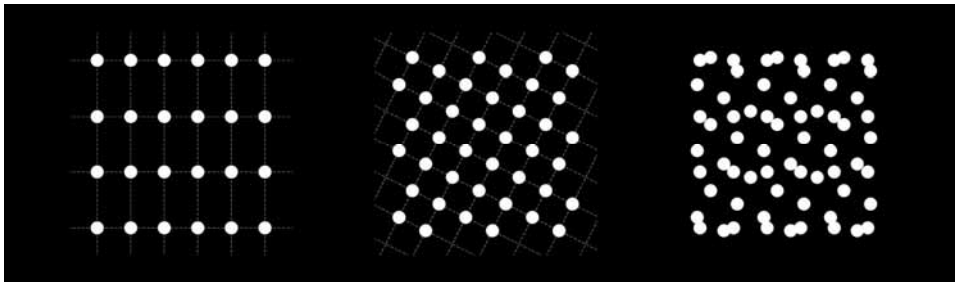


Figure 3: Example of decomposition method application: simple overlapping of two regular meshes leads to a much more complex form

is to highlight homogeneity of the form based on unification of its parts and their complete subordination to the general geometrical pattern of that form. Measure A is presented in a chart (fig. 1). An example of using measure B is the simulation presented in fig. 8.

2.3 Dispersion

The main feature of the method is using a random (stochastic) factor in the process of shaping the form. It is also necessary to distinguish a number of equal constituent parts. Their large number enhances the dispersion effect. Building of a form involves independent transformation of those constituent part (e.g. mov-

ing, rotating, changing proportions, color, etc.). Although only one and the same type of transformation is applied to all constituent parts, the way it is used (for each part) depends on a random factor. The method can be used for rebuilding the initial form of an orderly structure. By using dispersion, the structure undergoes gradual transformation losing its original clarity. This principle is presented in fig. 1. The use of a random factor can be measurable as a pre-set simulation parameter. Similarity of dispersed forms is an interesting issue. It turns out that two forms can be similar, although none of pairs of their constituent parts is identical. This has been presented in the simulation (fig. 4).



Figure 4: Similarity of dispersed forms

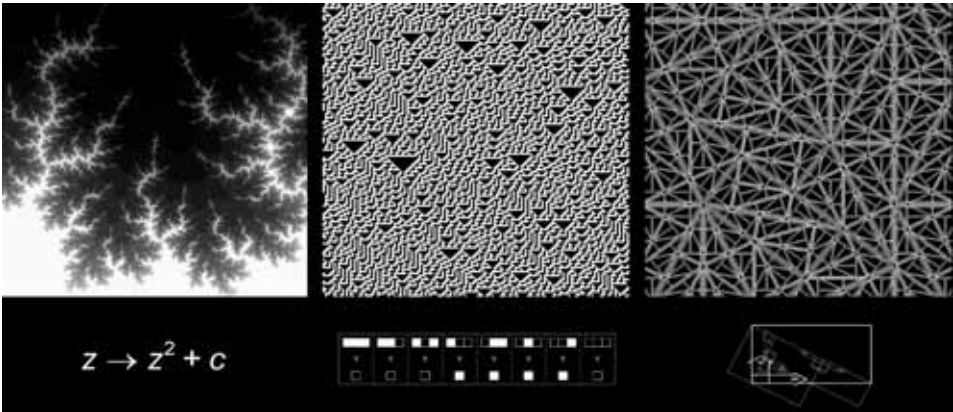


Figure 5: The method of new complexity: complex forms (upper part) are defined by simple rules (lower part): M set, 1D CA and IFS simulations

2.4 New complexity

The method aims at shaping complex forms by a simple process of formation. The principle of organizing (building) such forms is described as a system of higher order [15, 16]. The method refers to the mathematical theory of deterministic chaos, fractal geometry, and emergency. Rules can be expressed in different ways, for instance according to mathematical, fractal and emergent models. Each of the models defines a separate category of form definition. In the mathematical model, the form is defined using various formulas and rules. Computer simulations enable to observe basic systems of higher order at a purely theoretical level, e.g. Mandelbrot set (fig. 5a). In the emergent model, the process involves solely determining spatial relations and rules of mutual influence between particular elements of the form. An example is a simulation of cellular automaton CA (fig. 5b) [19]. The principle of the fractal model is based on transformations that can be described in the form of a simple geometrical pattern. An example of the above is a simulation using IFS [11] (fig. 5c).

3. COMPLEXITY IN ARCHITECTURAL CREATION

The chapter presents possible applications of decomposition, deformation, dispersion and new complexity methods for interpreting contemporary architecture facilities (since 1980s to present).

3.1 Decomposition in architecture

Deconstructivism is a clear example of the effort towards developing complex forms in architecture by using the method of decomposition. It started in the early 1980s, and matured in 1988 when the Museum of Modern Art in New York (MoMA) organized and exhibition of 'Deconstructivist Architecture'. The exhibition presented works by the then little known architects, such as Bernard Tschumi, Rem Koolhaas, Peter Eisenman, Daniel Libeskind, Zaha Hadid, Frank O. Gehry, Helmut Swiczin-

sky and Wolf D. Prix. The prevailing trait of works by deconstructivists is heterogeneity of architectural forms. In other words, they use a number of independent and intentionally unsuitable parts and divisions which do not result from a logical structure, as well as clashes and mergers of various elements. A rooftop remodeling above Falkestrasse in Vienna, designed by the Coop Himmelb(l)au Group in 1983 (fig. 6) is a vivid example of the above with its small scale and significance for a general architectural arrangement. On the one hand, Jacques Derrida's philosophy had a major influence on the development of the trend. On the other, geometry is the language of deconstruction. This can be seen in designs by Bernard Tschumi (in Parc de la Villette: clash between orthogonal networks of pavilions and natural landscape and deconstruction of cube) and Peter Eisenman (use of geometrical diagram). Deconstructivism continued developing until the end of the 20th c. as an expression of vanguard. Although it has finished, it still influences the contemporary architectural thought being a source of cautious or not cautious inspiration.

3.2 Deformation in architecture

The endeavor of contemporary architects aimed at larger complexity of forms also include designs focusing on fluid, flexible and streamlined forms. Instead of straight lines, various curves are used. Planes are replaced by complicated non-ruled surfaces. In the middle of 1990s, folding stood in opposition to deconstructivism in the contemporary architecture. In 1993, a special edition of *Architectural Design* was entitled 'Folding in architecture'. The editor was architect Greg Lynn. The article by the editor signified a new turn in architecture, and one chapter of the article was particularly telling: 'Curving away from Deconstructivism' [8]. The new trend developed under the influence of philosophical works by Gilles Deleuze. On the other hand, attempts made by architects to preserve homogeneity of form and combine various elements of the composition are deeply

rooted in geometry and mathematical topology. Möbius strip and Klein bottle are frequently referred symbols. Contemporary architects try to blur boundaries between notions considered to stand in opposition to one another, such as: interior and exterior of a building (e.g. Möbius House, designed by Ben van Berkel), or material and virtual space (e.g. Virtual Trading Floor, designed by Asymptote).

The idea of folding fits into presuppositions of the deformation method presented in the article. New computer modelling tools (e.g. NURBS) and new building techniques (CNC prefabrication) became catalysts for the architectural development in the field. A pioneer example of using new technologies is the Guggenheim Museum in Bilbao by Frank O. Gehry. Below you can see photographs of the museum with computer simulations by the author resulting from topological transformation of three cubes (fig. 7). Application of the deformation method also includes the blob architecture. An example of the above is Peter Cook and Colin Fournier's Kunsthaus in Graz (fig. 8).

3.3 Dispersion in architecture

Although the dispersion method is used for designing many contemporary architectural facilities, it is hard to define any wider ideological background for it, specific trend or style which would highlight the significance of stochasticity in shaping architectural forms. The issue was a subject of few research projects [e.g. 6, 17]. Example of using the methods is the Holocaust Monument in Berlin by Peter Eisenman. Within an orderly orthogonal system of 2711 concrete blocks we may distinguish irregularities (slight deviations from vertical plane and variable height of elements). Those irregularities give a specific aesthetic dimension of the whole facility. Dispersion is associated with natural conditions, perception of which is closer to a man. We may compare, for example, a wall made of various size stones, free surface made of granite brick, and tecton-

ics of medieval tenement houses with similar forms of regular and repeated structure. The majority of method applications refers to the composition of facades. Randomness can be fixed in a building (e.g. distribution of windows and colors of materials in Sharp Centre for Design in Toronto by William Alsop) and may result from the way those facilities are used (e.g. mobile light breakers in National Library in Paris by Dominique Perrault). Yet another example of the application of a random factor is the Torre Agbar building in Barcelona by Jean Nouvel. Its external concrete shell is perforated by 4400 random square openings. Additional layers are placed on the main construction, including corrugated steel sheet cladding of variable colors (from red to blue) and glass panels as light breakers (fig. 9).

3.4 New complexity in architecture

The development of forms according to the new complexity method is something new in architecture, and searching for possible applications in its early stage. However, inspirations derived from the mathematical theory of deterministic chaos and fractal geometry already appeared in the output of several architects as well as publications by architecture critics [5] in 1990s. For example, the design of a new opera house in Cardiff of 1994, American architect Greg Lynn used a method which he described as 'branching'. Branching referred to fractal geometry [7]. However, there is a difference between inspiration and application. We are able to provide only a few examples of architectural facilities developed, where high complexity results from a deterministic process. It is also important that the specific nature of designing such processes and forms changes our understanding of creation in architecture.

One of preliminary examples of new complexity being applied in architecture is a multi-function complex of buildings at Federation Square in Melbourne designed by Australian architects Donald Bates and Peter Davidson in



Figure 6: Decomposition method in architecture: left: rooftop remodeling above Falkestrasse in Vienna (by Coop Himmelb(l)au, 1983), right: project development of the University in Frankfurt am Main (by P. Eisenman, 1987)



Figure 7: Guggenheim Museum in Bilbao (by F. Gehry, 1997) in comparison to computer simulations made by author – as an example of deformation of architectural form



Figure 8: Computer modeling of blob forms – example of deformation. Left: simulation with use of Bryce 5.5 program; right: Kunsthaus in Gratz (by P. Cook, C. Fournier, 2003)



Figure 9: Application of dispersal in the composition of the facade of office building – Torre Agbar in Barcelona (by J. Nouvel, 2006)

2003. The composition of all facades is determined by 'pinwheel tiling', an non-periodic tiling developed and described several years earlier by American mathematician Charles Radin [13]. A simple and repeatable construction produces a complex result (fig. 10). Although the complex is a rather direct quotation from mathematics, the mechanism of developing its form can be planned. The rule of 'pinwheel tiling' can be described using the IFS method [15, 11]. Minor modifications of the rule comprise a specific encyclopaedia of complex forms, and intuitive predicting of the result extends beyond human imagination (fig. 10). Another later example of a facility is the centre of water sports in Beijing developed for the Olympic Games of 2008 (by PTW Architects), where the geometrical form results from using the Weaire-Phelan model.

4. CITY COMPLEXITY AND URBAN PLANNING

Methods of complex form classification presented in the article can also be used in urban planning as tools for interpreting the development of a city and its structures. However, it is not possible to extrapolate architectural observations directly on urban planning. It results from basic differences between the notion of architectural and urban creation. The goal of architecture is to develop building facilities from scratch. The immediate spatial context, neighbourhood of other buildings, can be important for design decisions. However, by nature, creating architectural facilities is a one off activity enclosed in time (from design to building). Obviously, the process of city development is much more complicated and has a different nature. Usually, a city already exists at the moment of 'urban intervention'. It is developed continuously. Thus, complexity is not a predetermined goal but a starting point for designing. For these reasons, contrary to architectural creation, an analytical process is a basis for urban planning.

4.1 Application of the complexity classification in urban planning

Methods for developing complex forms: deconstruction, deformation, dispersion and new complexity have been formulated to support methodical division and creation of architectural forms. In urban planning, the method gain their new meaning and other applications.

The decomposition method in architecture is closely linked with ideas of deconstruction, and can be a tool for various geometrical experiments and computer simulations, which may inspire the process of architectural design. In the urban scale, an equivalent of applying the method is imposing a new structure in a city, such as rebuilding of Paris in the 19th c. according to plans by Georges Eugene Haussmann. The decomposition method can be a mere tool for a methodical organization of such spatial transformations of a city. The potential of the method in creating new geometrical simulations in urban planning is doubtful (may lead to controversies).

The deformation method in architecture is mainly expressed by effort of creating fluid, liquid forms and using complex ruled or non-ruled structures. However, both presuppositions of the method and origin of architectural folding are much broader. The first pioneering examples of the trend is an urban project of the Rebstockpark Housing Estate in Frankfurt on Main by Peter Eisenman (1990) [15]. The urban parametric modeling (UPM) [18, 10] fits into the formula of the method, which proves the possibility of recording and transforming the structure of the city as a set of combined elements from the scale of projection to façade detail.

The dispersion method in architecture, which introduces a random factor to the composition, has the strongest influence on the façade or outer surface of a building. The significance of stochasticity in urban planning is much broader, naturally embedded into the process of planning and the structure of the city. Master plans usually define objectives at a general level while reserving space for individual

architectural solutions (e.g. color, composition of facade, roofing, variable height of a building within permissible scope, etc.). Possible applications of the dispersion method in urban planning are much broader than in architecture. New tools can be used for e.g. analyzing plans and examining the morphology of urban structure.

New complexity, as a method of creating or interpreting forms, is more developed in urban planning than in architecture. It is sufficient to refer to research by Michael Batty presented in the several publications and books: 'Fractal Cities' (1994) [2] and 'Cities and Complexity' (2007) [1]. Using emergence in urban planning is expressed in methods of simulating urban growth using cellular automata (CA) or simulations supported by the agent based modeling. The scope of searching for the system of higher order focuses on the area of analysis rather than creating new spatial solutions. However, nowadays the potential resulting from advanced urban analyses is crucial for the direction of city development.

4.2 The nature of the city complexity

What is the source of complexity in contemporary cities? Is it the result of stochastic activities or phenomena (related to dispersion method)? Is the organization based on mechanisms described in the article as a part of the new complexity method? Trying to find answers to those questions is inspiring in itself. Although one cannot expect equivocal results, each examining the boundaries between various categories of complexity of the urban tissue broadens our knowledge about the structure of a city and enhances our designing capacity. This issue has been a subject of various research [1, 12]. A fractal organization can be found in some simple urban systems. The language of the new complexity method can be used to describe an African village of Ba-ila [4], or plans of renaissance ideal cities. Figure 11 below presents examples of computer simulations developed using IFS and cellular auto-

mata CA.

The degree of complexity of contemporary cities is much larger and even if we find specific mechanisms of mathematical self-organization, undoubtedly the stochastic factor is an important component of their structure. Figure 12 includes an example of using the dispersion method while analyzing a part of Berlin, Germany. Consecutive stages of the simulation use simple transformation of the city structure (shifting, rotating). Slight 'movement' of buildings in Berlin changes completely its urban tissue. Differences in urban systems blur at various stages of the simulation, which can define the degree of the primary organization. An attempt to interpret complexity was made by the author in his exhibition of 'Images of Complexity' [14] which artistic aim was to look for relations between the structure of a city and dispersion and new complexity (fig. 13).

A visual complexity of a city provides a separate plane for interpreting. Urban structure reflected in a projection is a mere simplification. In fact, a man perceives a city as a set of thousands of various views, combined into a certain whole in minds of inhabitants [9]. Tall buildings have particular significance for developing of the 'image of a city', which range of visual impact is in principle larger [3, 20]. The development of simulation techniques creates possibility of precise analysis which is a subject of research under the 2TaLL Project. Further research should focus among others on defining comparative relations between visual complexity and complexity of the city geometrical structure.

5. CONCLUSIONS

Tendency of shift from classic understanding of systematics towards creation of complex forms in contemporary architecture is clearly seen. It is influenced by new techniques of computer modeling and digital CNC prefabrication. In recent years we can also observe indisputable progress in sphere of modeling and

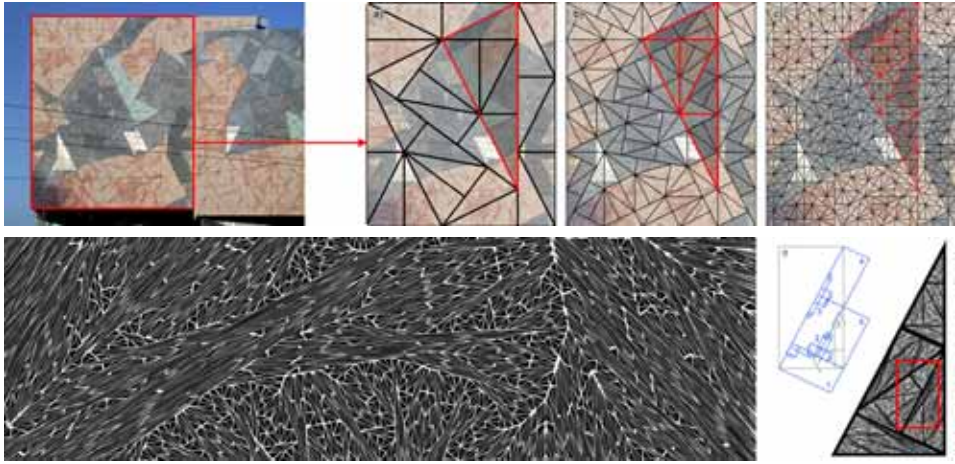


Figure 10:Composition of the facade of Federation Square in Melbourne (by Lab A-S) – example of application of new complexity in architecture. Original view of façade (above) and example of radical reorganization of composition by micro modification of its construction scheme

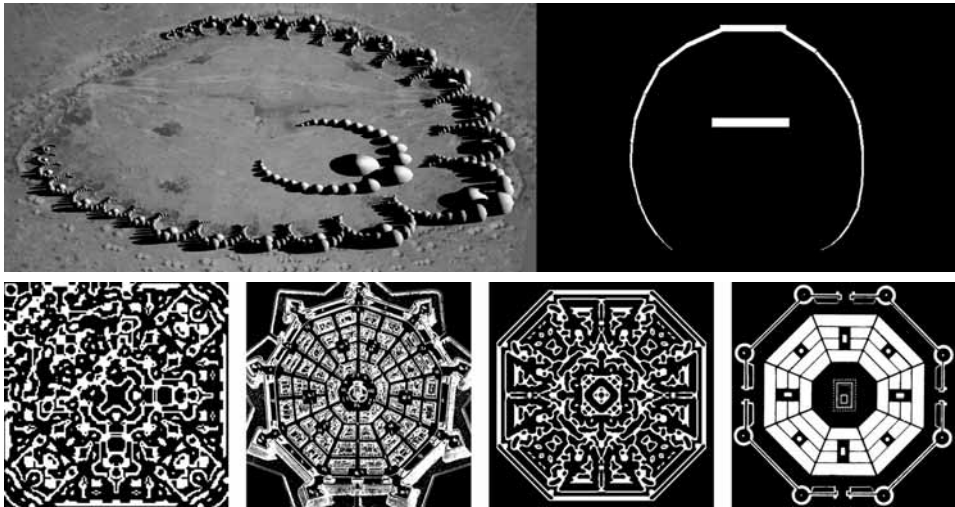


Figure 11: Sample computer simulations prepared with application of IFS method and cellular automata (CA) method. Above: African village of Ba-ila, below: plans of Renaissance ideal cities compared to simulations of cellular automata (CA)

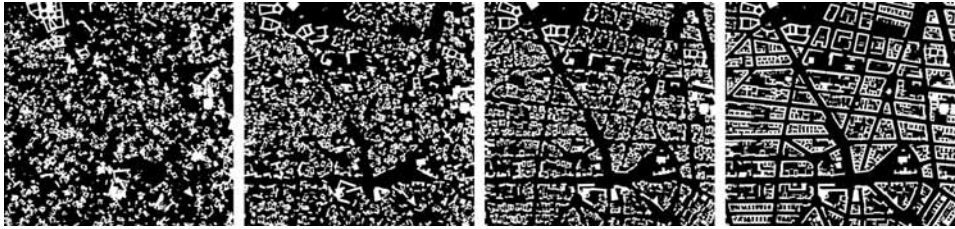


Figure 12: Dispersion method used in analyzing a part of Berlin, Germany



Figure 13: Interpretation of city complexity as a part of authors exhibition 'Images of Complexity' (Toruń 2014)

visualizations of cities. Quantity and accuracy of accessible 3D virtual city models increases rapidly. However, in architectural design the purpose is creation of new forms (buildings), so in urban planning the key importance is process of analysis of city. 3D city models allow application of advanced computer simulations for exploring complex urban structures. Both in architecture and urban planning, important subject of creation or subject of analysis is complexity. Geometry becomes universal language necessary for creation of bases for new digital techniques and also for better interpretation of structure of complex forms. The article presents proposal for systematics of complex forms within the four methods: decomposition, deformation, dispersion and new complexity. General examples of application of each method in architecture and urban planning are presented as well. The proposed methods allow better understanding of construction of complex forms. They can be applied as both, interpretation tool for architectural developments, or urban structures. Each or the methods can be a basis for development of specific computer simulations possible to encode in geometry.

Therefore, the presented methods can be a design tool as well.

ACKNOWLEDGMENTS

This research was funded by a Norwegian Financing Mechanism. Digital model of Berlin provided by Berlin Partner GmbH. All other 3D city models used for simulation were made by Cyber Urban Center at WPUT Szczecin. I gratefully acknowledge this support.

REFERENCES

- [1] M. Batty. *Cities and Complexity*, The MIT Press, Cambridge, Massachusetts, 2005.
- [2] M. Batty and P. Longley. *Fractal Cities*, Academic Press Limited, San Diego, 1994.
- [3] K. Czyńska. Tall Buildings and Harmonious City Landscape. In *Space & Form* 2010/13, pages 267-280, Szczecin, 2010.
- [4] R. Eglash. *African Fractals: Modern Computing and Indigenous Design*, Rutgers University Press, 1999.
- [5] Ch. Jencks. *The New Paradigm in Architecture. The Language of Post-Modernism*, Yale University Press, 2002.

- [6] R. Krawczyk. The Codewriting Workbook, Princeton Architectural Press, 2008.
- [7] G. Lynn. Animate Form, Princeton Architectural Press, NY, 1999.
- [8] G. Lynn. Architectural Curvilinearity: The Folded, the Pliant and the Supple. In *Folding in Architecture*, pages 24-31, John Wiley & Sons, 2004, (first published in: *Architectural Design* 8/1993).
- [9] E. Morello and C. Ratti. A digital image of the city: 3D isovists in Lynch's urban analysis, In: *Environment and Planning B: Planning and Design* 36(5), 837-853, London, 2009.
- [10] Z. Paszkowski and P. Rubinowicz. Toward the Parametric Modeling in Architecture and Design. In *Proceedings of 7th ICECGDG*, pages 33-36. Kraków, 1996.
- [11] H. Peitgen, H. Juergens and D. Saupe, *Fractals for the Classroom. Part 1: Introduction to Fractals and Chaos*, Springer, NY, 1992.
- [12] J. Portugali. *Self-Organization and the City*, Springer, Berlin, 2000.
- [13] Ch. Radin. *Miles of Tiles*, American Mathematical Society, 1999.
- [14] P. Rubinowicz. *Images of Complexity*, exhibition in Wozowania Art Gallery, Toruń, 2014.
- [15] P. Rubinowicz. *Chaos as the Higher Order in Selected Trends of Contemporary Architecture*, doctoral thesis, Cracow University of Technology, 2011.
- [16] P. Rubinowicz. *Chaos and Geometric Order in Architecture and Design*. In *Journal for Geometry and Graphics*, Volume 4, Number 2, pages 197-207, Heldermann Verlag, Lemgo, 2000.
- [17] P. Rubinowicz. *Parametric Modeling – Random Factors in Architecture*. In *Proceedings of 8th ICECGDG*, pages 81-85, Austin, 1998.
- [18] N. Steinø. *Developing a Parametric Urban Design Tool: Some Structural Challenges and Possible Ways to Overcome Them*. In *Architecturae et Artibus* 2014/1, Białystok, 2014.
- [19] S. Wolfram. *A New Kind of Science*, Wofram Media Inc., 2002.
- [20] A. Zwoliński. *A day in a shadow of high-rise*. In *Architecturae et Artibus* 2014/1, pages 67-71, Białystok, 2014.

ABOUT THE AUTHOR

Paweł Rubinowicz, PhD architect, assistant professor at West Pomeranian University of Technology Szczecin, Poland. Participant of 7th, 8th ICECGDG and 9th ICGG conferences. Doctoral thesis about application of deterministic chaos theory in contemporary architecture. Since 2009 curator of the Architects Gallery Form in Szczecin. Since 2013 member of Cyber Urban Center research team at WPUT. Contact: pawel@rubinowicz.com.pl

EYESIGHT CARTOGRAPHIES – UNFOLDING THE VISUAL SPHERE

José Vítor CORREIA, Manuel Couceiro da COSTA, Ana Santos GUERREIRO and Luís ROMÃO

University of Lisbon, Portugal

ABSTRACT: This paper addresses the relationships between the acknowledgment of different perspective concepts and its application in free-hand conceptual and observational drawings. A cartographical interpretation of perspective is pursued, by departing from the notion of “visual sphere” – a virtual spherical surface surrounding a viewer and upon which the entire set of visual data is laid. This work followed the developments of a R&D project where a computational tool was created aiming the reinforcement of the role of perspective in architectural design. The base for that research was the Extended Perspective System (EPS), a new geometrical perspective concept which was translated to a mathematical formulation and then written into a computer algorithm. After the development of the computational implementation, a role in the drawing learning process was envisaged and outlined in the form of a basic didactical strategy [1]. Therefore, we proceeded to create a non mandatory course – suggestively named Eyesight Cartographies – where that strategy could be tested. Cylindrical and spherical perspective methods were introduced and practiced as complementary paradigms, regarding the prevalent linear perspective paradigm. Also, recognizing that underneath each of these perspective systems is a single cartographic procedure, i.e., a particular method to map the visual sphere, students were challenged to scrutinize and inquire on the visual significance of several other methods found in the cartography realm, when used for a perspective depiction purpose. We introduced theory in a first moment and developed a workshop in a second. The spectrum of the course goes from the practice of free-hand drawing to the writing of computational algorithms dedicated to the automation of perspective drawing.

Keywords: Perspective, curvilinear perspective, visual perception, free-hand drawing, extended perspective system, perspectograph.

1. INTRODUCTION

The learning and practice of free-hand perspective drawing in architectural design education promotes the awareness of the place of the self in relation to the space and its visual and physical experimentation. By the exploration of the reversible variations of depth and spatial immersion, the process of perspective drawing consequently also conveys an awareness of the architectural phenomenon and constitutes an important and critical means to clarify its visibility. Besides the features of direction, relative sizes and positioning (conceptually, attributes that altogether inform and perform representa-

tion, identification and construction of referents in visibility and spatiality), perspective symbolizes the experience of depth, by assigning distance and hence the convening of the fourth dimension: time. The linear perspective code, although having established an important criterion of visual verisimilitude, intrinsically restricts the reference to a panoramic and kinetic spatial experience. The overcome of this paradoxical condition requires an inquiry into the conceptualization of the classical perspective. To rethink and re-equate the perspective representation, by congregating the linear and the curvilinear perspectives in a hybrid conceptualization, seems to fit the contemporary con-

sciousness of spatial fluidity, acknowledging the motional and visual dynamics of the observer. Consequently, the perspective representation should no longer be just static, with a motionless observer symbolized in the hierarchy established with the horizon line and the symmetrical vanishing points, as elements that bring order to the representation.

This poses new challenges to the didactics of drawing in architectural courses, particularly, by requiring an increased understanding of the conceptuality that is inherent to geometry and of the formal abstractions which are the base of architectural design. However, hand drawing remains an important asset, as a training from the visible and an effective “representational epiphany” for the architectural becoming.

2. FROM THE “REAL” TO THE “WIDE”

In the development of the perspective paradigm there is a decisive moment, whose importance is widely recognized imposing perspective as the epistemological model of classical occidental thought. It corresponds to the period of the discovery of the *costruzione legittima*, in the 15th century, with a special role from Brunelleschi, on account of his merit for the recognition of the identity between Point of View and Principal Vanishing Point. In this period the geometrical perspective codes were established and responding to a fundamental question which was the “Real” representation, the mimesis, overcoming older intuitive and naive experiments in the field of perspective, such as the fish bone attempts (Figure 1).

Since then, the evolutionary process of representational drawings, until then imprisoned by that fundamental question, got free and new paradigms emerged, allowing for technical and conceptual drawings, scenography, *trompe l’oeil*, anamorphosis, stereotomy, axonometry, descriptive geometry, curvilinear perspective, etc. In the last example - curvilinear perspective - the corresponding fundamental question is now the “Wide” picture, i.e. the representation not of what the eye sees in a fixed direction but moreover the agglutination of a set of sequential views into one single drawing, al-

lowing the understanding, exploration and representation of space in a completely new way. The power and consequences of this new regard are not yet fully understood, but there is a huge potential in several distinct fields as arts, architectural epistemology, esthetics, philosophy, history, virtual reality, among others. The development of the paradigm of the “Wide” has been slowly and also crossed trial and error, by adopting diverse curved projection surfaces, whose lines of drawing must somehow be transferred to the picture plane. We have intuitive examples of the graphical strategy of bending lines dating from ages ago, as exemplified through a drawing (Figure 2) by Jean Fouquet (1640), but the more scientific approaches started in the 20th century, being developed through the work of M.C. Escher, with his studies and establishment of geometric cylindrical structural grids [2] for his drawings. Also, the work of Barre and Flocon, building the geometric codes for spherical perspective in the already classical book *La Perspective Curviligne* [3] or the work of Gérard Michel, connecting the intuitive approach of the urban sketchers [4] with some pragmatic geometrical rules for the delineation of cylindrical and spherical perspectives [5]. But the potential of bending lines goes beyond planar, cylindrical and spherical perspective, being possible to explore hybrid stages of the projection surface and its transfer into the representation surface. That is the overall capability of a new conceptual instrument - the Extended Perspective System [6], which congregates linear and curvilinear perspectives, from principles to practice and from automation to didactics, so fully exploring the “Wide” paradigm.



Figure 1: Fish bone structure, in a decorative mural fragment, 1st century, [7].



Figure 2: Cylindrical perspective attempt, “The arrival of Emperor Charles IV at Saint Denis, [8].

3. AN APPROACH FOR A NEW SYNTHESIS

Linear perspective, with its convenient picture plane, is a well known and vastly applied representational system. Along with its historical weight and broadness in modern visual culture, it really renders objects in a visually congruent manner, mainly by depicting all lines as graphic straight lines. Significantly, its conception and operativity are aimed at narrow fields of view, much below 180° .

Cylindrical perspective, by the use of a cylindrical projection surface (a simple curvature surface), allows an up to 360° panoramic delineation of the horizontal space surrounding the observer, but also restricts the field of view, vertically, to a value much below 180° . Generically, lines are turned into sinusoidal curves, with the exception of vertical lines, which remain straight.

Spherical perspective, by the use of a spherical projection surface (a double curvature surface), conveys a omnidirectional consistent view of the space surrounding the observer, only refrained by the progressive anamorphic character of the depiction when going towards the rear area of the global visual field. Nonetheless, this system of perspective has noticeable figurative capability till a 270° field of view. Generically, lines are turned into transcendental curves, with the exception of lines parallel to the visual axis, which remain straight.

The rectilinear graphical operativity of linear perspective has much practicability in free-hand drawing, by appealing to the common sense of a straight line. Nonetheless, that representational system is designed to convey

spatial information that is mainly in front of the observer, within a restricted field of view. On the other hand, the curvilinear systems allow the translation of a sight in motion, targeting a broader or even complete field of view, which constitutes an important supplementary feature, regarding linear perspective capabilities. But those systems require the graphical management of mathematically specified curved lines, which turns their operativity more intricate and therefore less used in the didactics of free-hand drawing. Most of all, the rectilinear and curvilinear approaches to perspective stand historically as alternative or even rivaling propositions, which restrains their overall complementary action as referentials to drawing.

In order to sustain a more inclusive and versatile method for the didactics of perspective in free-hand drawing, we took into account the formulation of the EPS, a broad theoretical concept of perspective that has been firstly presented [9]. Since 2010, this concept was embraced by a multidisciplinary academic team, gathering skills in architecture, drawing, geometry, mathematics and computation, in a R&D project called NAADIR [10]. The main objective of the project was the making of a computational implementation of the EPS, with the purpose of turning this representational system usable in architectural design, from conception stages to final presentation drawing. The computational approach was crucial to the operativeness of the EPS, since this concept stands mainly as a dynamic curvilinear model of perspective, strongly dependent on the mathematical calculations that make feasible its intrinsically wide range of possible perspective depictions. A more detailed description of the EPS and the corresponding software tool that was created is already published [11].

Fundamentally, the EPS is established upon three principles. The first, inherited from the curvilinear perspectives, is the dissociation of the projection surface and the representation surface (picture plane). The projection surface is the surface where primary projections of 3D points are set upon, while the representation

surface is the planar surface where those projections are then transferred onto, in order to obtain the final perspective depiction to be visualized. The second principle is the mutability of the projection surface. Instead of the single static projection surface that is ahead in the formulation of the referential perspective systems, the EPS determines the use of a spheroidal parametric projection surface. By means of homological procedures, this surface can be turned into the linear, cylindrical or spherical projection surfaces that are inherent to the former perspective systems but, principally, it can assume any transitory spheroidal state between those boundary configurations. Any particular state is defined by the combined numerical concretization of two parameters: Radius and Eccentricity. Therefore, this mutable projection surface continuously fills the void between the autonomous and separated geometric foundations of the spherical, cylindrical and linear perspective systems. The third principle is the adaptability of the method for transferring the projected information from the projection surface to the representation surface. In linear perspective, the method is direct: projections become the depiction elements. In cylindrical perspective, the method is natural: the unrolling of the developable projection surface. In spherical perspective, the method is protocolled: a selected cartographic projection to render depiction elements onto the representation surface. In the EPS, since the projection surface is mutable, consequently the transferring method is reconfigured dynamically, adjusting either to each of the boundary states or to each of the transitory states of that surface, aiming visually optimized results. The Figure 3 shows an array of the diverse EPS hybrid depictions that can be obtained by variations of the parameters Radius (R) and Eccentricity (E).

So, in general terms, the EPS is a unified concept of perspective representation, as it congregates and articulates the linear, cylindrical and spherical perspectives, as referential boundary systems, in a single theoretical build. The departing point for the formulation of the

EPS was indeed the acknowledgment of the specific representational capabilities of each of those perspective systems, but mainly the envisioning of a truly complementary role to be fulfilled by them, as long as some particular changes in their overall geometric foundations take place. A role in the didactics of drawing can also be envisaged from here.

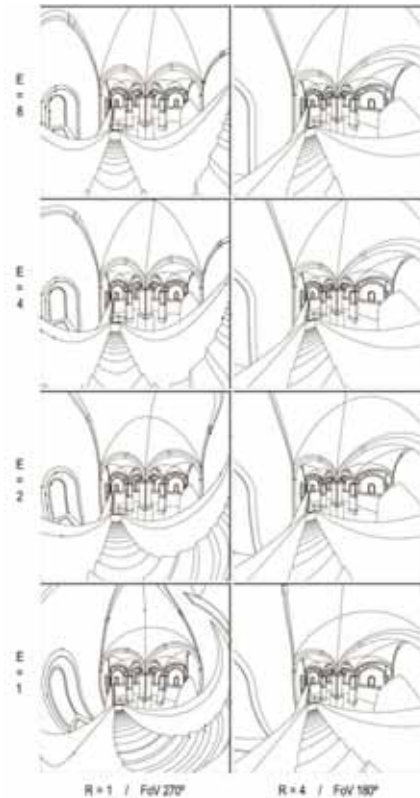


Figure 3: An array of images resulting from variations of the EPS parameters.

4. A NEW DIDACTICAL STRATEGY FOR DRAWING

Although linear perspective shares many attributes that are identifiable in direct visual

perception, it does not constitute an innate response in free-hand drawing. On the contrary, to draw what and how one sees is undeniably a difficult task to accomplish, without proper training and learning. Particularly, among architecture students this difficulty is sensed and requires much attention from drawing and geometry teachers. The didactics of linear perspective concepts and methods has a significant role in this process. In fact, it most frequently allows students to firstly get an objective comprehension and awareness of the vision process itself, before even the issue of drawing from it is tackled.

Traditionally, geometry and drawing disciplines follow different training methodologies. In geometry classes, priority is given to projection concepts and to its effects in the delineation of geometric figures. Here, standard exercises such as drafting cubes in space with 1, 2 and 3 main vanishing points (Figure 4) effectively help students surpassing the natural naive stages of perspective drawing and to recognize the importance of a geometrically structured approach to representation. These exercises, in a unrestricted field of view version (Figure 5), also emphasize and turn visible the conventional condition of linear perspective delineations - a code that, although rectilinear, may render less close to vision graphical representations. Counterpointing linear perspective, a supplementary acknowledgement of the available representational resources can be obtained by drafting in the cylindrical and spherical perspective systems, eventually with the help of underlying grids to overcome the usual difficulties regarding curved lines (Figure 6).



Figure 4: Perspective drawings with 1, 2 and 3 vanishing points, within a regular FoV

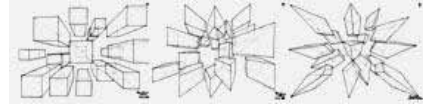


Figure 5: Perspective drawings with 1, 2 and 3 vanishing points, within a wide FoV

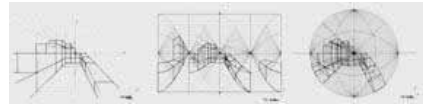


Figure 6: Drawings in linear, cylindrical and spherical perspectives, with the aid of lattices

In drawing classes, a more heuristic approach is pursued, by favoring an intensive training in observational drawing from reality. It is there expected that individual visual intuition and reasoning takes place, triggering students' improvement and differentiation in free-hand drawing skills.

In a wider context, even the observational drawing, between urban and architectural spaces, as well as landscapes and gardens, increase insights for the subjective comprehension of the representation scale in a free-hand drawing. There are some situations where the students understand that the linear perspective is not enough to dominate some space representations. This gap requires responsive versatility, also between free-hand representation scales and their relation with the drawing support likewise contributes to the intuitive research about other graphic solutions (curvilinear perspectives) more inclusive (and even more immersive) beyond the linear perspective paradigm, now no more longer used here as an undeniable graphic feature. And these resources, which become skills, are an added value for the versatility of graphic resolutions for designing conceptions, important domain in project of architecture.

It is from the side of the results achieved in drawing lessons that we get the more interesting clues on the management of linearity when it comes to the visual approach to representa-

tion. In fact, within the NAADIR project tasks proceeding, it was possible to collect several sets of observational drawings made by students under controlled circumstances, in order to reduce variables and turn their subsequent analysis more consistent. Generally, it was identified a correlation between the growth of students skills in drawing and a flexible use of the graphical rules of linear perspective. In fact, while students have shown the capability to pursue thoroughly an explicit linear perspective lattice, that structural support becomes more implicit, many times even diffuse, when previously acquired geometric knowledge is confronted with the real time visual data that feeds the mind/gesture process of drawing. Plus, when students are asked to draw a surrounding architectural space, by allowing the sight to move freely in a wide field of view, targeting up, down, left and right in the scene, the linearity of the drawings noticeably deviates towards an implicit curvilinear lattice, although many times a scattered one. Intuitively, apprentices of drawing seek a graphical response to the experience of a dynamic observation of space (Figure 7). Significantly, the bending of the lines in these drawings is many times readable as a spontaneous deduction of the cylindrical and spherical graphical configurations of perspective, with their curved horizontal and vertical delineations.

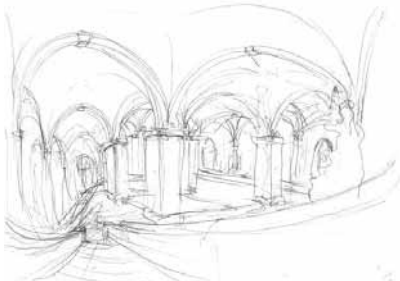


Figure 7: A student's observational drawing of the Mãe-de-Água building in Lisbon

These remarks lead us to interrogate on the sufficiency of the current perspective model as basis for the didactics of drawing, particularly

in architectural courses where it is crucial that students become rapidly fluent in the use of the drawing resources to assist the learning and practice of architectural conception. We see that the acknowledgment of linear perspective concept and method constitutes a gravitational reference for the embodiment of a structured and consistent way of drawing. In practice, this role means that a free-hand drawing's structure orbits that referential paradigm, getting closer or further from it as visual priorities and judgment take place. We also see that enhanced drawings seem to deviate from a structural linear perspective grid towards a diffuse perspective model that somehow resembles the alternative curvilinear systems. Therefore, we foresee a potential role of the spherical and cylindrical perspective systems as also gravitational references to be included in a more holistic perspective model for the didactics of drawing. This would be the outcome of the EPS concept in this realm.

Alongside with the much valuable learning of the classic linear perspective system, students would additionally be trained in the extreme concepts of cylindrical and spherical perspectives and in over all possible hybridizations sustain by the EPS concept. Demonstrations and trial with the EPS computation tool should also provide an enriched set of possible perspective depictions. The expected result of this more diverse and extensive training would be the settlement of an amplified referential framework for the individual process of growth and mastering in free-hand drawing. We believe that, within this didactical strategy, an apprentice shall find more efficiently his own personalized hybrid way(s) of drawing. Therefore, we proceeded to create a non mandatory course - suggestively named Eyesight Cartographies - where that strategy could be tested.

5. EYESIGHT CARTOGRAPHIES - A DIDACTICAL EXPERIMENT

The Eyesight Cartographies course promoted the implementation of our multidisciplinary research. In a first experimental implementa-

tion it was addressed to college students of the third level of architecture, this course gathered some subjects: geometry and drawing, graphic and programming computer, mapping and cartography.

The first two classes, predominantly expositive, presented the first point of the program: tridimensional perspectives: origins and formulations. The perspectival renaissance paradigm and its evolution was seen and commented on images of Art History and Visual Communication, from Renaissance to the present. Examples of *trompe l'oeil*, anamorphosis, curved frames, bicentral projections, stereoscopy, convergence and binocular vision, reading relief and depth were given. The students were asked to do, as homework, three inclusive and intuitive free-hand drawings from their workroom, from different points of view that the representation of that area was the most broad and complete of it as possible. These preliminary representations will converge as an introduction to the need to understand the ever new applications of perspectival representation and its use, both symbolic and symptomatic of the contemporary.

After the first moment, where the historical evolution of the "wide" paradigm of perspective was presented and its practice intuitively approached, the didactic structure of the course followed through the learning, comparison/synthesis and deeper understanding of the "triangular" boundary of the hybrid possible states of the new perspective images. The "triangular" boundary has, as vertices, the geometrical structures of planar, cylindrical and spherical perspectives. Assuming that the first vertex was a knowledge pre-requisite of the students, emphasis was given to cylindrical and spherical structures, namely considering direction (vanishing points and lines), dimension/proportion and space location control, then practical experienced as free-hand drawings.

After the first two modules of the course, where theory and practice were directed to analogue drawing, the third and last module was dedicated to the automation of drawing

with digital means, namely by the use of computer graphics procedures. At this stage, geometrical, mathematical and programming knowledge were evoked and gathered in order to configure and produce new perspective depictions alternative to the conventional linear perspective system. To this purpose, a cartographical approach to perspective was pursued by establishing two main principles: 1st, the concept of 'visual sphere' - a virtual spherical surface surrounding a viewer and upon which the entire set of visual data is laid; 2nd, the notion of perspective as the result of transferring that set of visual data to a planar surface - a map - where the spatial representation can then be seen as a perspective depiction.

The concept of visual sphere can be inferred from ancient optics studies and constitutes a way to rationalize geometrically the experience of a complete visual field, i.e. the whole set of visual data collected by an observer that fully rotates his eyes and head. Also, the enlarged concept of perspective sustained by the prior consideration of the visual sphere turns the perspective mechanics similar to the cartographical mechanics (considering the whole globe cartography). At this point, linear perspective and cylindrical and spherical perspectives must be pointed out as just particular results of the application of specific cartographical methods, namely the gnomonic projection, the cylindrical perspective projection and the azimuthal equidistant projection. It must be emphasized that these three methods are among a group where approximately two hundred methods are currently catalogued. Therefore, a research on the use of several other methods to perspectival purposes seems to be pertinent, at least has an hypothesis investigation. So, students were challenged to scrutinize and inquire on the visual significance of several other methods found in the cartography realm. Cartographic criteria of classification - equidistance, equivalence and conformality - become also relevant in perspective, as they determine the ability of a specific method to preserve the visual magnitudes along lines, the visual pro-

portions of elements in the visual field and the local visual configuration of shapes. Also, other features of cartographic projections were considered, such as the ability to preserve curvature continuity (despite curvilinearity) or the ability to generate a continuous mapping of the visual field, without gaps in the visual data.

For this first edition of the Eyesight Cartographies course, a set of twelve representative methods were then selected, taking into account the classified groups of cartography, shown at the next

The student research work was made in groups and attended four stages. The first stage was the building and validation of a 3D digital architectural model to constitute the subject of the perspectives to be generated. The second stage was the computational implementation of the projection of the 3D model in the surface of the visual sphere. From this moment on, this spherical projection was taken as the primary data for the calculations to be made in order to implement the cartographical method within each work group. The third stage was common to all the groups and regarded the computational implementation of the azimuthal gnomonic projection that leads to linear perspective. This stage was important for students to get a cartographic perception of the conventional method of perspective and also to get an understanding of some basic procedures of computer graphics needed to tackle the research issue. Finally, the fourth stage was the overall computational implementation of the specific cartographic method that was chosen. The work involved the acknowledgement of the particular geometrical and /or analytical procedures, the algebraic translation and the writing of the computer program itself. After the implementation is achieved and the graphical delineations become feasible, the groups made a preliminary analysis of the visual properties of the perspective system that was configured. Besides intrinsic properties, each new system was compared to the prevalent linear perspective system, emphasizing differences and complementary qualities.

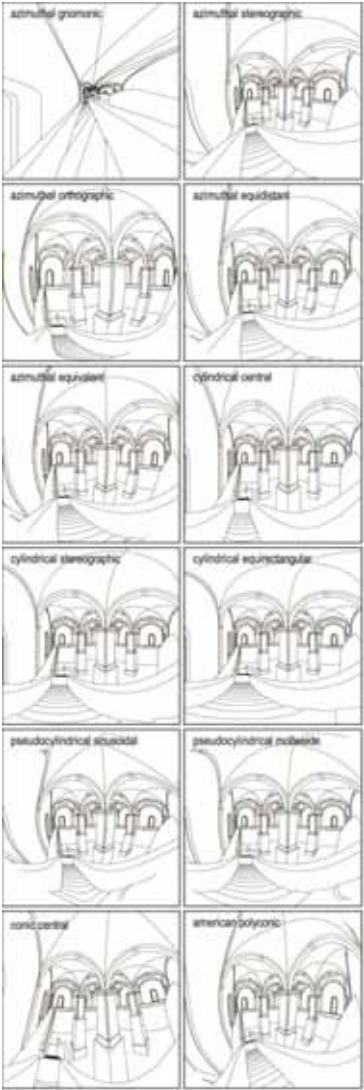


Figure 8: Perspective delineations with twelve different cartographic projections

6. CONCLUSIONS

This paper argues that the proficiency on free-hand perspective drawing is one of the fundamental objectives that an architecture student must pursue, in order to gradually let

drafting be the easy and natural way to support and reveal the visual and spatial reasoning that is inherent to architectural conception. In this learning process, drawing and geometry disciplines together provide complementary methodologies, by the meeting of applied visual intuition with the rationale of perspective graphical codes. In its inherent plasticity, although, enhanced free-hand perspective drawings tend to develop graphical interpretations of visual perception that do not entirely depend on the observance of the graphical rules of a single perspective system, namely linear perspective. Instead, driven by real time visual judgment, they seem to gain an autonomy that strongly relies on a flexible use of the geometric principles of perspective. A new formulation of perspective - the Extended Perspective System - develops a holistic approach by congregating and hybridizing linear and curvilinear perspectives, therefore building up an expanded and upgraded concept of perspective. After the development of a computational implementation, that firstly sustained the operativeness of this new perspective system, a role in the drawing learning process was envisaged and outlined in the form of a basic strategy, where the main concepts of the Extended Perspective system would inspire a renovated didactics of drawing in architecture graduation. This motivation took the form of a new discipline called Eyesight Cartographies, a multidisciplinary approach to perspective drawing that intends to provide students with a comprehensive understanding and critical attitude towards the standards of space graphical representation.

The student's interest and commitment in the Eyesight Cartographies course, as well as the results achieved, reinforced this didactical approach as a pertinent added value in the current teaching of drawing in the architecture graduation. As a major conclusion from this experiment, we understand that, in the next editions of the course, after the computational module, a return to the freehand drawing practice should be promoted.

We believe, given the good results of this

first didactic and pedagogical application of our research on the increment of this procedural hybridization, both for the field of design either for learning. For greater freedom and know-how acquired for re-interpret or redraw the reality in architectural and graphic-spatial representation.

Corroborating these facts, it should be noted the very recent inclusion of this course in the curriculum of the doctoral program in Architecture, of the Faculty of Architecture of the University of Lisbon-.

ACKNOWLEDGMENTS

The authors would like to acknowledge the students that did the course and produced some of the graphic work shown in this paper.

REFERENCES

- [1] Correia, and al. Bending Lines – Mastering the linear world of Perspective. *M&D2013, Tucuman. Argentina*, 2013.
- [2] B. Ernst, *Le Miroir Magique de M.C.Escher*, Fribourg (Switzerland), Medea Diffusion S.A., 1987.
- [3] A. Barre and A. Flocon. *La Perspective Curviligne*, Paris, Flammarion. 1984.
- [4] E. Salavisa et al. () *Urban Sketchers in Lisbon – drawing the city*, Lisbon, Urban Sketchers, Quimera Editores, 2012.
- [5] G. Michel, L'oeil au centre de la sphere visuelle, Porto, *Boletim da Aproved* n°30/2013, pp.3-14. 2013.
- [6] J. Correia and L. Romão. Extended Perspective System, Predicting the Future [25th eCAADe Conference Proceedings / ISBN 978-0-9541183-6-5] Frankfurt am Main (Germany), September, pp. 185-192, 2007.
- [7] E. Panofsky. *Perspective as Symbolic Form*, Zone books, NY. 1997.
- [8] A. Barre and A. Flocon. *La Perspective Curviligne*, Paris, Flammarion, 1984.
- [9] J. Correia and L. Romão. Extended Perspective System, Predicting the Future [25th eCAADe Conference Proceedings / ISBN

978-0-9541183-6-5] Frankfurt am Main (Germany), September, pp. 185-192, 2007.

[10]J. V. Correia et al. A New Extend Perspective System for Architectural Drawings, *Global Design and Local Materialization*, [15th International Conference, CAAD Futures 2013, Shangai, China, July, ISBN 978-3-642-38973-3], pp. 63-75, 2013.

[11]J. V. Correia et al. A New Extend Perspective System for Architectural Drawings, *Global Design and Local Materialization*, [15th International Conference, CAAD Futures 2013, Shangai, China, July, ISBN 978-3-642-38973-3], pp. 63-75, 2013.

ABOUT THE AUTHORS

1. José Vitor Correia is an architect, graduated in 1993 by the Faculty of Architecture of the Technical University of Lisbon, where he teaches in Descriptive Geometry courses since then. Currently he is developing his doctoral thesis in the specialty of visual communication, within the architecture field. His research addresses the issue of geometrical Perspective, on its conceptions and on its use in architectural delineation.

2. Manuel Couceiro da Costa is a Professor at the Faculty of Architecture of TU Lisbon. He holds a Ph.D. in Visual Communication in FA.Ulissabon.

3. Ana Santos Guerreiro is a Professor at the Faculty of Architecture, Technical University of Lisbon. She holds a Ph.D. in Visual Communication by the same Faculty and a Post-degree in History of Art by the FCSH, Nova University of Lisbon. She is graduated in Fine Arts, by the Faculty of Fine Arts in Lisbon. She works and researches in Fine Arts and Theory of Art.

4. Luis Romão is an architect and a professor that holds a Ph.D. in architecture: design and computation, from the School of Architecture and Planning, MIT.

FIBONACCI TRIANGLES AND CIRCLE CHAINS AND A GOLDEN BI-ARC SPIRAL IN NON-EUCLIDEAN PLANES

Sybille MICK¹ and Gunter WEISS²

¹Graz University of Technology, Austria ²Vienna University of Technology, Austria

Dedicated to Professor Manfred Husty on the occasion of his 60th birthday.

ABSTRACT: The well-known Fibonacci numbers and the golden mean are dimensionless numbers, but are commonly visualized in the Euclidean plane by a nested set of squares and a nested set of golden rectangles having squares as gnomon figures. Quarter circles in the gnomon squares are connected to a smooth bi-arc spiral of a discrete spiral group. This spiral is called *the golden spiral*. Visualizations also can be done in general normed planes, affine planes, higher dimensional spaces, the latter again with a Euclidean structure. The geometric objects are nested sets of cubes and prisms or hypercubes and hyperprisms (cf. [13], [12]).

In this paper we present visualizations in the Euclidean plane that will work in a hyperbolic plane, where neither squares nor similarities exist as well as in a Minkowski plane – a normed (affine) plane with a centrally symmetric, simply closed convex curve as unit circle. The key idea is to replace the Euclidean Fibonacci squares by circles which are well-defined both in the hyperbolic plane and in Minkowski planes. We call a circle whose radius is a Fibonacci number Fibonacci circle and we construct a sequence of Fibonacci circles as follows: Three consecutive Fibonacci circles touch mutually externally. The third circle is chosen out of two possibilities such that it follows a “spiral” arrangement. We study properties of this arrangements and construct a smooth circular bi-arc “spiral”.

Keywords: Fibonacci sequence, golden mean, hyperbolic plane, Minkowski planes, golden triangle.

$$x^2 - x - 1 = 0. \quad (2)$$

1. INTRODUCTION

Historically, the golden mean was and still is visualized by a segment and a point dividing the segment into a major segment M and a minor segment m (Figure 1, [1], [2] and [11]) such that the ratio of their lengths is

$$|m| : |M| = \phi^{-1} = 0.618 \dots \quad (1)$$

The basic condition is

$$|m| : |M| = |M| : |M + m| ,$$

and as a conclusion the well-known fact that the ratio $|M| : |m| =: \phi$ fulfills the quadratic equation

The positive solution of (2) is

$$\phi = \frac{1+\sqrt{5}}{2} \quad (3)$$

known as *golden mean* that indicates the arithmetic mean between 1 and $\sqrt{5}$.

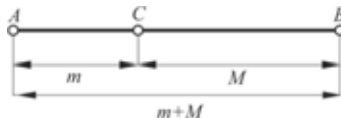


Figure 1: A one-dimensional visualization of the golden mean by three collinear points.

The Pythagoreans were the first to find this special ratio comparing the segment side and diagonal ratio of a regular pentagon. They recognized that these two segments are incommensurable, when applying the so-called exhaustion method, which, mathematically spoken, is the calculation of a continued fraction. Later on, the mathematical definition of the golden mean was and is based on the special continued fraction

$$\phi := 1 + \frac{1}{1 + \frac{1}{1 + \frac{1}{1 + \dots}}} = 1.618... \quad (3)$$

It also became more or less standard to use the symbol ϕ for this positive irrational number (cf. [1]). Replacing the leading number 1 in the denominators in (3) by another natural number gives the so-called metallic means [5]¹.

Remark: Dividing a segment justifies also the widely used concept *golden section* and *golden ratio*. The *ratio* is an affine invariant of three collinear and ordered points and it is a signed real number. The unordered set $\{A, B, C\}$ of three points allows the six sequential arrangements ABC, ACB, BAC, BCA, CAB, CBA on an oriented (affine) line. Thus, a geometric figure shows not only one number x , when visualizing a mathematical fact about ratios, but in some respects simultaneously the numbers

$$x^{-1}, 1 - x, \frac{1}{1-x} \text{ and } -\frac{1-x}{x}, -\frac{x}{1-x}. \quad (4)$$

Therefore, triples of points defining ratios with less than these six different (absolute) values are of geometric and mathematic interest. There are only two such triplets, c.f. [12]:

- a) ABC with C the midpoint of segment [A,B],
- b) ABC with C at the golden section of [A,B].

Visualizations of the golden ratio show not only to the single value $\phi = 1.618...$, but all the values (4).

2. FIBONACCI SQUARES AND THE GOLDEN SPIRAL

For visualizing the classical Fibonacci numbers $0, 1, 1, 2, 3, 5, 8, \dots, F_i, \dots$ it is common to use a set of squares with side lengths according to the F_i . The squares are arranged in a spiral manner such that, for $i \geq 6$, each new square A_{i+1} “touches” the triple of squares A_i, A_{i-2}, A_{i-3} (cf. [1]). Therefore, we can say that this construction starting with two adjacent unit squares works *from small to large*. At each stage the convex hull of the set of i squares is a rectangle with side ratio approaching towards $1:1.618...$ for $n \rightarrow \infty$, (Figure 2). Finite rectangles with this special side ratio are called *golden rectangles* (GR).

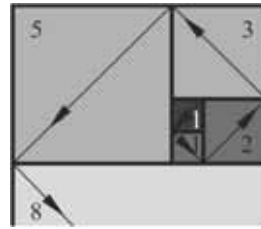


Figure 2: Fibonacci sequence of squares

On the other hand, visualizations of the golden mean (GM) commonly start with a finite GR and cut away a square as gnomon figure, leaving a GR as the rest. Continuing this cutting process delivers squares in a spiral arrangement, (Figure 3). Replacing the diagonals of the gnomon squares by quarter circles one gets *the golden spiral*, (Figure 3). The cutting process leads to a similarity factor $0.618... = \phi^{-1}$ and works *from large to small*.

¹ Periodic continued fractions and their values are studied e.g. in [5 - 8].

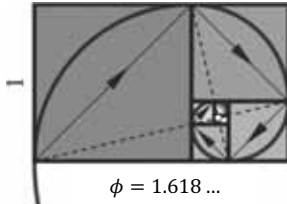


Fig. 3: Golden bi-arc spiral consisting of quarter circles

Fibonacci numbers F_i , Fibonacci quotients F_i/F_{i-1} and the golden mean value ϕ (and its invers ϕ^{-1}) are dimensionless numbers. Usually, visualizations are done in the Euclidean plane. But it is neither necessary to use a two-dimensional structure nor a metric structure. For example, any affine transformation of Figure 2 and Figure 3 would also deliver visualizations to these numbers. It is always possible to adjust an affine coordinate frame to an arbitrarily given parallelogram, such that it can be called a golden parallelogram (see Figure 4 and [4]). The arrow polygon in Figure 2 and Figure 3 is a one-dimensional visualization of the Fibonacci numbers, while the squares visualize a sequence of areas, namely $\{F_i^2, i = 1 \dots \infty\}$. As it is shown in [13] it is also possible to get 3D and higher-dimensional visualizations.

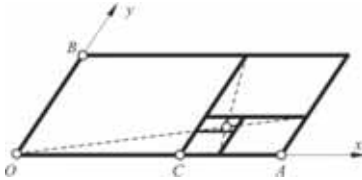


Figure 4: Affine coordinate frame to an arbitrarily given parallelogram such that it becomes a golden Parallelogram.

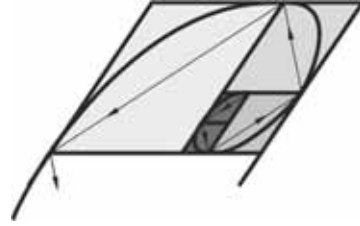


Figure 5: Spiral polygon as a one-dimensional visualization of Fibonacci numbers

3. FIBONACCI SEQUENCE AND GOLDEN MEAN VISUALIZED BY TANGENT CIRCLES

There are visualizations of the golden mean and the Fibonacci sequence in general metric planes, so-called Minkowski planes based on parallelograms, see [14]. In a hyperbolic plane \mathbb{H} , on the contrary, there exist neither similarities nor tilings with squares. Therefore, an attempt with squares and rectangles would not work. In order to find substitutes for the Euclidean squares we go for circles, as they exist in any Cayley-Klein plane and also in Minkowski planes (see e.g. [9] and [3]). We start with the Euclidean case.

3.1 Fibonacci numbers and golden triangles in the Euclidean plane.

To visualize the Fibonacci sequence 0,1,1,2,3,5,8,... and *golden spirals* we use the same principle as well in the Euclidean plane as in an arbitrarily given Minkowski plane or in the hyperbolic plane. We start with a point as null-circle c_0 and a pair of tangent “unit” circles c_1, c_2 containing this point. Next, circle c_3 touches c_2 and c_1 externally and its radius F_3 is the sum of the radii $F_2 + F_1$. If we choose any of the two options for c_3 , the centers O_1, O_2, O_3 of c_1, c_2, c_3 specify one rotational direction of the plane. Now, we are able to specify the recursion relation. For a triple of consecutive circles c_{i-1}, c_i, c_{i+1} the circle c_{i+1} touches c_i, c_{i-1} externally and has the radius $F_{i+1} = F_i + F_{i-1}$. In order to receive some sort of spiral polygon, we claim that the rotational

direction O_{i-1}, O_i, O_{i+1} equals the one of O_1, O_2, O_3 . We call the polygon with vertices O_j *Fibonacci sequence*. The centres O_j of three consecutive circles c_{i-1}, c_i, c_{i+1} form a triangle with side length ratios

$$\begin{aligned} (F_{i-1} + F_i) : (F_i + F_{i+1}) : (F_{i-1} + F_{i+1}) = \\ = F_{i+1} : F_{i+2} : (F_{i-1} + F_{i+1}). \end{aligned} \quad (5)$$

We call such triangles *Fibonacci triangles*.

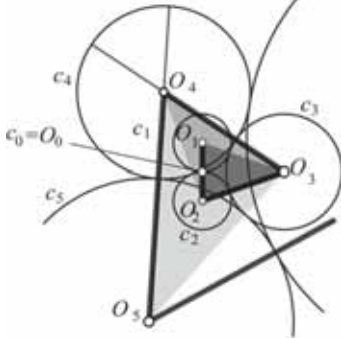


Figure 6: Fibonacci sequence visualized by a circle chain in the Euclidean plane.

In Figure 6 a Fibonacci sequence in the Euclidean plane is shown, whereby the positive rotational direction has been selected and maintained throughout the process of adding tangent circles.

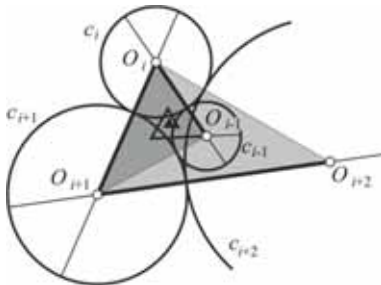


Figure 7: Sequence of tangent golden circles and golden triangles in the Euclidean plane.

The limit of the ratios (5) for $i \rightarrow \infty$ is

$$1 : \phi : \left(1 + \frac{1}{1+\phi}\right) = 1 : 1.618 \dots : 1.381 \dots \quad (6)$$

with ϕ indicating the golden mean. We call a triangle with such side length ratios a *golden triangle* GT, although in literature this term is used for isosceles triangles with side ratio $\phi : \phi : 1$ and $\phi : 1 : 1$, cf. [1]². On the other hand, we can start with three mutually tangent circles with radii $r_{i-1} = 1/\phi$, $r_i = 1$, $r_{i+1} = \phi$ in the Euclidean plane and proceed with the aforementioned construction. Then we get triples of consecutive circles with the same radius ratio. The sequence of centers O_j conveys a spiral polygon and all triples of consecutive centers are golden triangles GT.

Of course, it is possible to replace the sides of the polygon by circular arcs to receive a C^1 -continuous circular bi-arc spiral. In the Euclidean plane golden triangles might be of interest with respect to their remarkable points and objects.

3.2 Fibonacci numbers and golden triangles in Minkowski planes.

A Minkowski plane is a metric plane with a centrally symmetric, simply closed convex curve as *unit circle*, see [9].

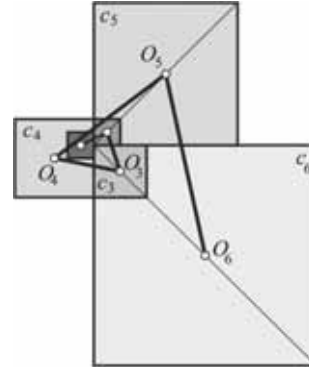


Figure 8: Fibonacci sequence visualized by a circle chain in a Minkowski plane with (Euclidean) squares as circles.

² In this paper we refer with *golden triangle* always to a non-isosceles triangle.

While in Euclidean planes we have translations, rotations, dilatations and symmetry, there are no (continuous) rotations in Minkowski planes. In general such a plane is not an inner product plane and orthogonality is a non-symmetric relation. Our first example is the Minkowski plane with the so-called Manhattan metric. The unit circle is a parallelogram and it can be visualized - without loss of generality - as a (Euclidean) square.

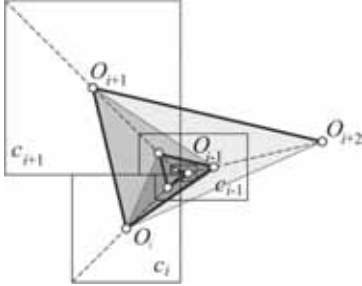


Figure 9: Minkowski-circle chain with radii progressing with the dilatation factor φ in a Minkowski plane with (Euclidean) squares as circles.

As our second example we present a Minkowski plane with homothetic (affine regular) hexagons as circles. They can be visualized as Euclidean regular hexagons.

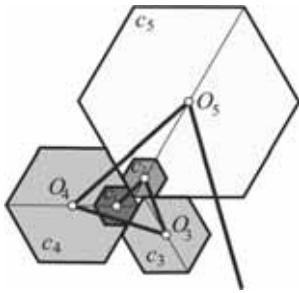


Figure 10: Fibonacci sequence visualized by a circle chain in a Minkowski plane with (affine regular) hexagons as circles.

In all Minkowski cases the centres of consecutive Fibonacci circles define a sequence of interlinked triangles with side ratios that converges to the side ratio of the golden triangle. For this reason we find a new sort of a *spiral polygon*.

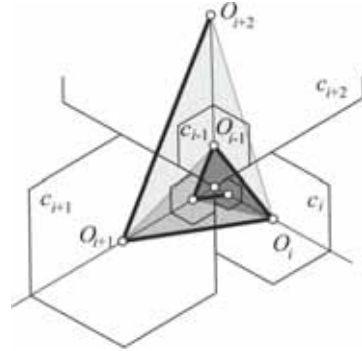


Figure 11: Sequence of tangent golden circles and golden triangles in a Minkowski plane with hexagons as circles.

The sides of this spiral polygon could be replaced by e.g. Minkowski-circular arcs. The replacement would not be smooth if the Minkowski-circles are not. In order to apply usual graphics software to perform the construction it might be an idea to use a curve c displayed as a Euclidean circle in the Minkowski plane. Figure 12 (right) illustrates the isometric image curve c' in the Euclidean plane of a curve c in the Minkowski plane with a square-shaped unit circle Figure 12 (left).

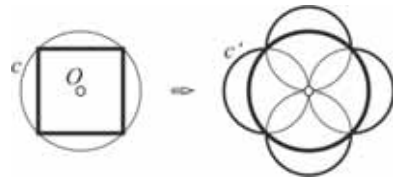


Figure 12: Image curve of c for the Minkowski plane with a square-shaped unit circle

In Figure 13 we present a “Euclidean looking” bi-arc spiral in a Minkowski plane with hexagons as circles.

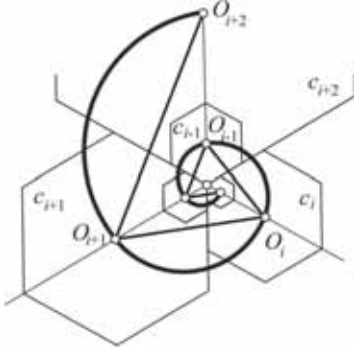


Figure 13: C^1 -Minkowski bi-arc “spiral” consisting of arcs according to Figure 12.

3.3 Fibonacci numbers and golden triangles in a hyperbolic plane.

In a hyperbolic plane we have reflections and rotations, but neither translations nor dilatations. But, like in Minkowski planes and the Euclidean plane (proper) lines have length ∞ . Starting from an arbitrarily given point on the line we can mark points with Fibonacci numbers as coordinates. Therefore, the approach of section 3.1 using a chain of tangent Fibonacci circles will also work in a hyperbolic plane.

For the hyperbolic plane we choose the Cayley-Klein model, the interior of an *absolute conic* \square embedded into the real projective plane. The construction of section 3.1 can be performed in the hyperbolic plane (see Figure 14). Again, we get a sequence of interlinked Fibonacci triangles, vertices are the centers O_{i-1}, O_i, O_{i+1} of consecutive Fibonacci circles. The sequence can be extended indefinitely by applying the recursion relation and the limit of the ratios (5) for $i \rightarrow \infty$ is (6).

Again we can construct a chain of circles, where the radii of consecutive circles have ratio $1:\phi$ (Figure 15). Triples of consecutive circles have centres O_{i-1}, O_i, O_{i+1} which are vertices of golden triangles and all the points O_i form

some kind of a spiral polygon in the hyperbolic plane.

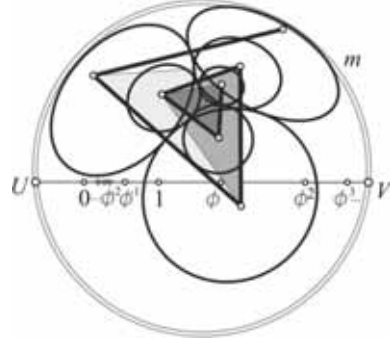


Figure 14: Fibonacci sequence visualized by a circle chain in the hyperbolic plane.

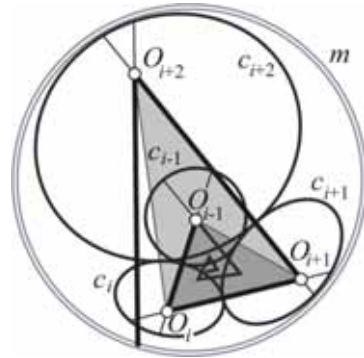


Figure 15: A sequence of tangent *golden circles* and *golden triangles* in the hyperbolic plane.

In the Euclidean case it is possible to replace the sides of the polygon by circular arcs such that we get a circular bi-arc spiral with C^1 -continuity and all arcs belong to the same center angle. In the hyperbolic plane we construct a circular C^1 -bi-arc to adjacent sides of a polygon (Figure 16). Due to the fact that golden triangles of different sizes have different angles, the construction delivers circular arcs of different center angles.

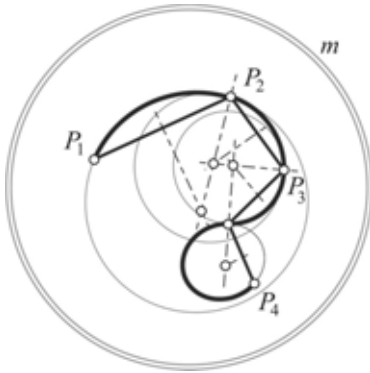


Figure 16: Circular bi-arc to adjacent sides of a polygon in the hyperbolic plane.

4. CONCLUSIONS

In this paper we present a uniform method to visualize the Fibonacci sequence and the golden ratio by sets of circles that are tangent to two preceding circles. The method works in the Euclidean plane as well as in Minkowski normed planes and the hyperbolic plane. Visualizing the golden ratio we get sequences of interlinked golden triangles such that two sides of these triangles belong to a spiral polygon. It is obvious how to proceed for Fibonacci type sequences belonging to Spinadel's metallic means m . There will occur sets of interlinked *metallic triangles* of side ratios

$$(7) \quad 1:m: \left(1 + \frac{1}{1+m}\right).$$

The set of these triangles for different values of m might be of interest on its own. In this paper we presented some Minkowski examples and the hyperbolic case. The application of the method to define *golden spiral polygons* in other Cayley-Klein planes will be the subject of an additional paper.

REFERENCES

- [1] A. Beutelspacher, B. Petri. *Der Goldene Schnitt*. 2nd ed., Spektrum-Akademischer Verlag 1997

- [2] R.A. Dunlap. *The Golden Ratio and Fibonacci Numbers*. World Scientific Publishing Co., 2003, ISBN 9810232640.
- [3] H. Martini, K. Swanepoel, G. Weiss. The geometry of Minkowski spaces - a survey, Part I, *Expositiones Mathematicae* 19 (2001) 97-142, ISSN 0723-0869
- [4] H. Martini, M. Spirova. Golden rectangles in normed planes. *Mitteilungen. d. Math. Ges. Hamburg* 29(2010), p.125-134 (MR 2012 a: 51019)
- [5] V.W. de Spinadel. Golden and Metallic Means in Modern Mathematics and Physics. *Proc. 13th ICGG 2008*, Dresden, Germany, ISBN 978-3-86780-042-6.
- [6] V.W. de Spinadel. Metallic Spirals. *Proc. 10th Int. Conf. APLIMAT 2011*, Bratislava, Slovakia, p.659-666.
- [7] V.W. de Spinadel, L.L. Echagüe. Espirales de Plata. *Journ. Math. & Design, Vol.8* (2008). ISSN 1515.7881
- [8] V.W. de Spinadel, A. Redondo Buitrago. Generalizing the Golden Spiral. *Journ. Math. & Design Vol. 11(1)*(2011), ISSN 1515.7881.
- [9] A.C. Thompson. *Minkowski Geometry*. Cambridge Univ. Press 1996, ISBN 0-521-40472-X
- [10] H. Walser. *Der Goldene Schnitt*. (Einblicke in die Wissenschaft) B.G. Teubner, Leipzig 1993, ISBN 3-8154-2070-9.
- [11] G. Weiss. Golden Hexagons. *Journal for Geometry and Graphics, Vol. 6* (2002), p.167-182.
- [12] G. Weiss: From the Golden Rectangle to the Laan Box and more. *Journ. Math. & Design Vol. 12* (2012), ISSN 1515.7881.
- [13] G. Weiss, V.W. de Spinadel. Bi-arc spirals in Minkowski planes. *Proc. 16th ICGG 2014*, Innsbruck, Austria, Art.Nr. 08, (to appear).

ABOUT THE AUTHORS

1. Sybille Mick, Dr. is a member of the institute of Geometry at the Graz University of

Technology (Austria). Her favourite fields of interest are non-Euclidean geometry and graphic education. She can be reached via e-mail: mick@tugraz.at

2. Gunter WEISS, Dr. is a retired professor for Geometry (since 2011) at the Dresden University of Technology, Germany. His e-mail address is weissgunter@hotmail.com .

Foldable Truss-Z module

Machi ZAWIDZKI and Takehiko NAGAKURA
Massachusetts Institute of Technology, USA

ABSTRACT: Truss-Z (TZ) is a modular skeletal system for creating free-form pedestrian ramps and ramp networks among any number of terminals in space. TZ structures are composed of four variations of a single basic unit. This paper introduces the concept of foldable Truss-Z module (fTZM) and presents the prototype scale model. The folding reduces the bounding volume of the module almost three times, specifically - by factor of 0.35.

Keywords: Foldable module, Deployable structure, Free-form, Truss-Z, Skeletal structure

1. INTRODUCTION

The TZ structures are composed of two units: R and L, which are mirror reflection of each other. Units are named according to the right-hand rule: the modules that ascend and “turn right” and ascend and “turn left” are called R and L, respectively. By rotation, the modules can be assembled in two additional ways (R_2 – rotated R and L_2 – rotated L), effectively giving four types of units. For further details see [1].

The problems of inexpensive fabrication of physical medium-size model of TZ modules have been presented at the 15th International Conference on Geometry and Graphics in 2012 in Montreal, Canada [2]. This paper introduces the concept of foldable Truss-Z module (fTZM), which substantially reduces its size for transportation or storage. The volume reduction ratio (VRR) compares the bounding volumes of the module in stowed (VB_s) and deployed (VB_d) states. VRR for fTZM is calculated as follows:

$$VRR[fTZM] = \frac{VB_s}{VB_d} = \frac{4.212 \text{ m}^3}{12.171 \text{ m}^3} \approx 0.35 \quad (1)$$

2. THE FOLDABLE MODULE

In principle, both Truss-Z modules (TZM) and foldable Truss-Z modules (fTZM) are based on a glide reflection in space. Figure 1 shows the exploded view of R-unit of fTZM and illustrates the plane of symmetry (reflection) and translation

vector. It also explains the naming convention for the module components.

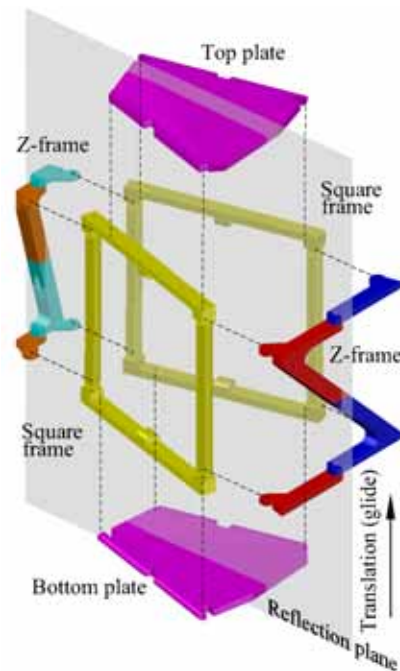


Figure 1: The symmetry and components of fTZM.

As shown in figure 1 the main component - square frame is reflected about a vertical plane and displaced along it vertically. The trapezoidal shape of the module in plan allows to create practically any path in the horizontal plane, while vertical displacement (glide) allows to move up (R & L) or down (R_2 & L_2). It was assumed that in the folding process the unit should collapse inside contracting all the external dimensions. This obviously posed additional constraints to the design. Glide reflection is also used for unfolding the unit, which is performed in two phases:

- Unfolding of the side Z-frames;
- Deployment of the bottom and top plates.

Figure 2 shows schematically both phases in respective columns.

2.1 Fabrication of the scale model

Making a physical model is an indispensable part of in this type of designing process. Here, the main consideration was the efficiency and practicability of folding, not the structural optimization. Therefore the elements were dimensioned intuitively, without proper structural analysis. As a result, the components of fTZM are most likely over-sized. For comparison and compatibility, the scale of the fTMZ prototype is the same as in [2], that is 1:32. The virtual model has been constructed in a CAD program. The actual fabrication using rapid prototyping methods was rather straightforward - the digital model has been 3D-printed and assembled with screws and bolts. Although making a functional virtual model was rather time-consuming, the fabrication process took only a couple of days.

Figure 3 illustrates the unfolding of fTZM prototype model. Each sub-figure corresponds to respective sub-figure in Figure 2.

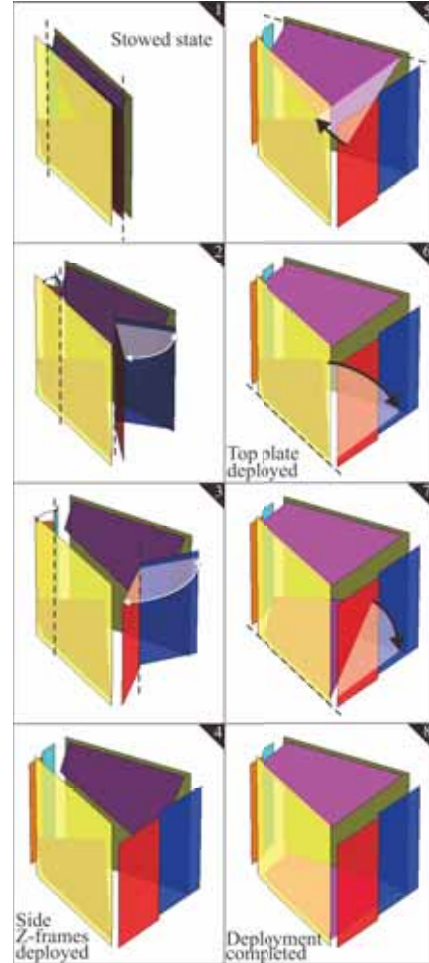


Figure 2: Left column: unfolding of the Z-frames. Right column: 5-6) Deployment of the top plate; 7-8) Deployment of the bottom plate. Colors correspond to Figure 1.

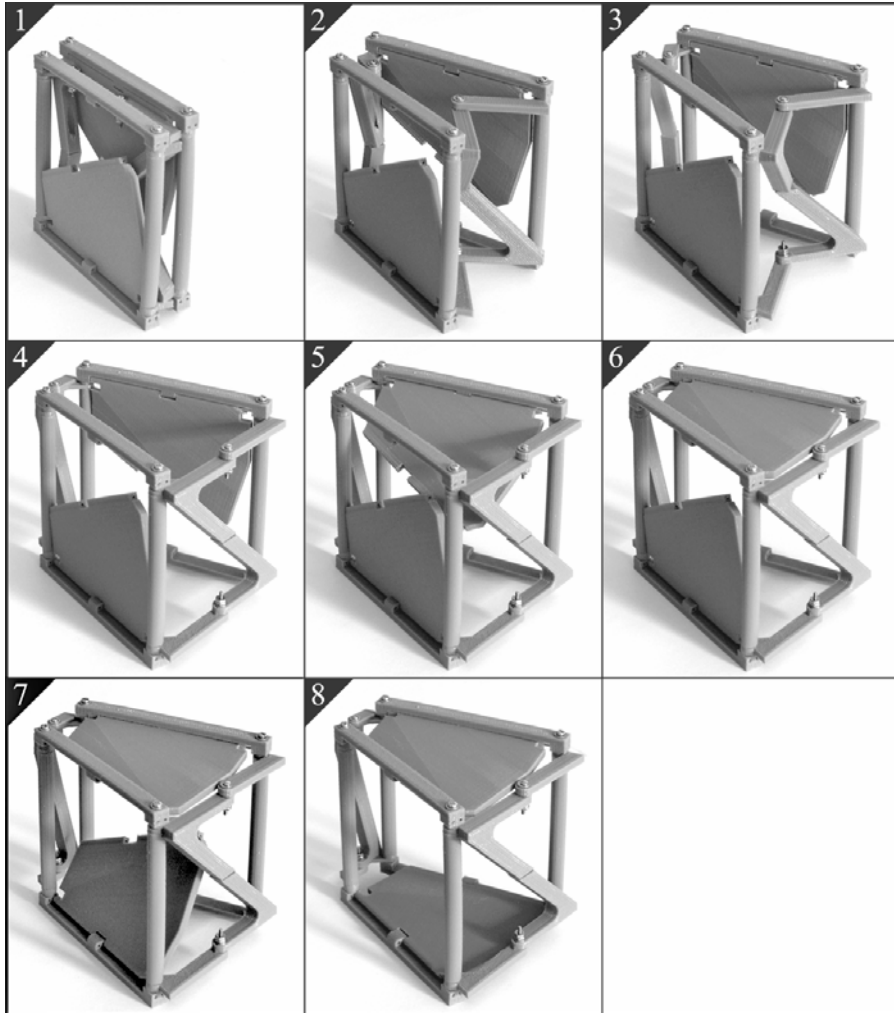


Figure 3: The unfolding of fTZM: 1) Stowed fTZM. 2-4) Unfolding of the side Z-frames; 5-6) Deployment of the top plate. 7-8) Deployment of the bottom plate.

For TZ system simple and efficient joinery between modules is crucial. In fTZM each subsequent unit “shares” the main square frame with the previous one, as shown in Figure 4.

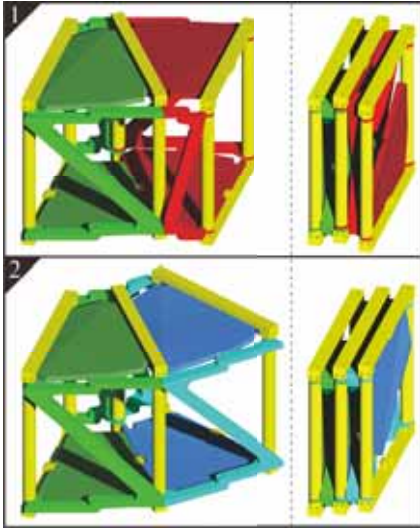


Figure 4: Two examples of two fTZMs assembled. On the left and right the deployed and stowed states are shown, respectively. 1) R and L units are shown in green and red, respectively. 2) Two subsequent R units are shown in green and cyan, respectively. The “shared” main square frames are shown in yellow.

3. CONCLUSIONS

Traditional deployable structures, although quite complicated as mechanical systems in their final form represent relatively simple, symmetrical geometries serving as: masts (linear), antennas (rotational symmetry), pantographic roofs (uniform grid), pool covers, domes, etc. This paper presents a system where the final assembly of foldable modules offers unprecedented geometrical freedom.

The main objective of this paper is the minimization of the volume reduction ratio (VRR).

In the solution proposed the subsequent modules “share” the main square frames, therefore it seems that connecting folded units would be difficult. Perhaps it would be more practical to assemble already folded units even at worsen compactness.

The module presented here needs a proper structural analysis and optimization. Moreover, there are a number of practical issues to be addressed, such as the deployment and locking mechanisms, placement of the supports, multiple-unit assembly etc.

ACKNOWLEDGMENTS

This work was completed as part of the Singapore University of Technology & Design and Massachusetts Institute of Technology Postdoctoral Fellows Program (SUTD-MIT PDP). The research is titled: *Effective computational methods for grid and raster-based modeling of practical problems in architectural and urban design layouts*.

REFERENCES

REFERENCES

- [1] M. Zawidzki and K. Nishinari. Modular Truss-Z system for self-supporting skeletal free-form pedestrian networks. *Advances in Engineering Software*, **47**(1), 147–159, 2012.
- [2] M. Zawidzki, K. Nishinari, T. Hanasaki and H. Sugimoto. Inexpensive fabrication of physical medium-size model of Truss-Z modules. *Proceedings for ICGG 2012: the 15th International Conference on Geometry & Graphics*, Montreal, Canada, August 1–5, 2012.

ABOUT THE AUTHORS

1. Machi Zawidzki, D.Eng., M.Arch. Dr. Machi Zawidzki is currently a postdoctoral fellow at the Department of Architecture of Massachusetts Institute of Technology, funded by the Singapore University of Technology and Design Postdoctoral Program.

The research is titled: “Effective computational methods for grid and raster-based modeling of practical problems in architectural and urban layout design”. In 2013 he has completed a postdoctoral research at the Research Center for Advanced Science and Technology of the University of Tokyo, funded by a grant from the Japan Society for the Promotion of Science. The research was titled: “Improvements of the Seniors’ Quality of Life through Application of Innovative Computational Systems to the problems of Accessibility, Ergonomics and Housing & Living Environment”. He received Eng. and M.Arch degrees from Warsaw University of Technology, Poland, in 2000 and the D.Eng degree from Ritsumeikan University, Japan, in 2010. From 2002 to 2005 he worked at Howard Alan Architects, Chicago, a studio specializing in sustainable and in particular - solar architecture. From 2006 to 2007, he worked at Studio E Architects, London, England, a studio specializing in sustainable and public architecture.

2. Takehiko Nagakura, PhD. Professor Nagakura is an architect from Tokyo. At MIT, he teaches courses related computer-aided design, and his research focuses on the representation and computation of architectural space and formal design knowledge. He has founded and led Architecture, Representation and Computation group (ARC) since 1996. His completed building, Gushikawa Orchid Center in Okinawa, Japan has been awarded with SD Review Award (1998) and Nikkei Kyushyu District New Office Award (1999). He is the co-author of *Gendai Kenchiku no Hassou (Ideas in Contemporary Architecture)* (Maruzen, 1989) and translated William J. Mitchell’s *the Logic of Architecture* (MIT Press, 1990) into Japanese (Kajima, 1991). His essays include “Shape Recognition and Transformation” in the *Electronic Design Studio*, edited

by William J. Mitchell, Patrick Purcell and Malcolm McCullough (MIT Press, 1990). The result of his past MIT studio courses in electronic design environment is summarized in a co-authored article, “Digital Pinup Board – The Story of the Virtual Village Project” in *Virtual Design Studio*, edited by Jerzy Wojtowicz (Hong Kong Press, 1995). Before coming to MIT in 1993, he worked for Fumihiko Maki in Tokyo, and was an instructor at Harvard University, Graduate School of Design. He earned Bachelor of Engineering in Architecture from Tokyo University in 1985, Master of Architecture from Harvard University in 1987, Master of Engineering in Architecture from Tokyo University in 1988, and completed his PhD at Harvard in 1996. In 1985 he received the prestigious Ishizaka Memorial Foundation scholarship from the Japanese Federation of Economic Institutions. He is the recipient of the Japan Information Culture Society Grand Prize in 1999.

FORMING THE SPLINE SIMILAR RULED SURFACE STRIP

Denis Sergeevich KORCHAGIN and Konstantin Leonidovich PANCHUK

Omsk State Technical University, Russia

ABSTRACT: In theory ruled surfaces methods of formation the ruled surface strips are known. For example in the work "Computational Geometry Line" by H. Pottmann and J. Wallner [1] for this purpose interpolation and approximation algorithms are used. In our work for building the ruled and non-ruled surface strips the dynamic characteristics of geometrical figures and solids, known in theoretical mechanics are used.

Keywords: The center of masses, the inertia ellipsoid, the directing line, ruled and non-ruled surface strips.

1. INTRODUCTION

Geometrical figures and bodies possess dynamic characteristics. In particular, they possess the axial and centrifugal moments of inertia and the centers of masses.

The axial moments of inertia are calculated by formulas known in theoretical mechanics [2]:

$$\begin{aligned} J_x &= \int (y^2 + z^2) \sqrt{(dx/dt)^2 + (dy/dt)^2 + (dz/dt)^2} dt; \\ J_y &= \int (x^2 + z^2) \sqrt{(dx/dt)^2 + (dy/dt)^2 + (dz/dt)^2} dt; \\ J_z &= \int (x^2 + y^2) \sqrt{(dx/dt)^2 + (dy/dt)^2 + (dz/dt)^2} dt, \end{aligned} \quad (1)$$

where $x = x(t)$, $y = y(t)$, $z = z(t)$ - parametric equations of a line.

Centrifugal moments of inertia are also evaluated by the well-known formulas [2]:

$$\begin{aligned} J_{xy} &= \int xy \sqrt{(dx/dt)^2 + (dy/dt)^2 + (dz/dt)^2} dt; \\ J_{yz} &= \int yz \sqrt{(dx/dt)^2 + (dy/dt)^2 + (dz/dt)^2} dt; \\ J_{xz} &= \int xz \sqrt{(dx/dt)^2 + (dy/dt)^2 + (dz/dt)^2} dt. \end{aligned}$$

$$\begin{aligned} J_{yz} &= \int yz \sqrt{(dx/dt)^2 + (dy/dt)^2 + (dz/dt)^2} dt; \\ J_{xz} &= \int xz \sqrt{(dx/dt)^2 + (dy/dt)^2 + (dz/dt)^2} dt. \end{aligned} \quad (2)$$

The axial and centrifugal moments of inertia form an inertia tensor. The tensor of inertia is a matrix of a square form [2]

$$J = \begin{bmatrix} J_x & J_{xy} & J_{xz} \\ J_{xy} & J_y & J_{yz} \\ J_{xz} & J_{yz} & J_z \end{bmatrix}. \quad (3)$$

Its elements enter into the equation of a surface of the second order

$$\begin{aligned} J_x x^2 + J_y y^2 + J_z z^2 + 2J_{xy} xy + 2J_{yz} yz + 2J_{xz} xz &= 1. \end{aligned} \quad (4)$$

This surface is an ellipsoid of inertia.

Based on the described dynamic characteristics and corresponding ellipsoids of inertia, we suggest method the forming of surfaces.

2. ELEMENTS OF THEORY

Consider the method of forming the spline similar ruled surface strip. We offer to form a fragment of ruled surface strip in accordance with the logic of the algorithm of formation of

a cubic spline. Below the fragment of ruled surface strip we will understand the set of ruled sections joined together on the common segments of straight line generators (generate segments). The ruled sections are the components of ruled surface strip limited by generate segments. Discrete carcass is the initial data in our offered method and it consists of generate segments. Thus, the strip is formed by continuous moving of a generate segment through beforehand discretely specified generate segments. Moving the generate segment of a fragment of ruled surface strip is produced along a directing line. This line is connected continuously with the specified generate segments. It contains the centers of masses of generate segments of a projected fragment of a ruled surface strip. Each generate segment of a continuous search carcass is considered in its own local coordinate system attached to the directing line. The directing line is calculated in particular cases so that each of n generate segments of an initial discrete carcass was given in the local coordinate systems. These local coordinate systems are n relevant Freinet trihedrons of this line [3].

The described dynamic characteristics and corresponding ellipsoids of inertia are considered together for the following reason: they help to assign the local coordinate system for each of the specified generate lines, which is associated with the directing line. It meets the requirements to the projected directing line.

Note that the directing line is the first phase and a necessary condition for the implementation of process of forming of a ruled surface strip. The second phase is the surface coverage continuous carcass lines of generate segments defined by the directing line.

2.1 Formation of the directing line

Consider the process of formation the directing line. The process contains the algorithm which repeated for each generator of a specified discrete carcass. In general the generators can be spatial (closed or not closed) curves. They can be specified in the basic Cartesian coordinate

system $Oxyz$ by the equations (for example in a parametric form $x = x(t)$, $y = y(t)$, $z = z(t)$), or by a discrete set of points. The disposition of the generators can be arbitrary with respect to the base coordinate system.

Assume the density of each generator of a constant value. Then the center of mass of the generator represents the point. The position of this point characterizes the distribution of points of the generator along its length. The center of mass is a geometric point invariable relative to the generator. Coordinates of the center of mass O_c of the generator determined by the well-known equations in the base coordinate system [4]:

$$\begin{aligned}\bar{x} &= \frac{1}{L} \int x \sqrt{(dx/dt)^2 + (dy/dt)^2 + (dz/dt)^2} dt; \\ \bar{y} &= \frac{1}{L} \int y \sqrt{(dx/dt)^2 + (dy/dt)^2 + (dz/dt)^2} dt; \\ \bar{z} &= \frac{1}{L} \int z \sqrt{(dx/dt)^2 + (dy/dt)^2 + (dz/dt)^2} dt,\end{aligned}\quad (5)$$

where

$$L = \int_{t_{\min}}^{t_{\max}} \sqrt{(dx/dt)^2 + (dy/dt)^2 + (dz/dt)^2} dt$$

is the length of generator.

Following step after determining the center of mass of the generator will be the transition from the base coordinate system to the local coordinate system combined with the centre of masses of the generator. This transition is made by parallel transfer the axes of the coordinate system $Oxyz$ from point O to the point O_c (the center of mass of the generator). As a result of the transition obtain the equations $x_c = x_c(t)$, $y_c = y_c(t)$, $z_c = z_c(t)$ defining the generator in the coordinate system $O_c x_c y_c z_c$.

Next we turn to the ellipsoid of inertia. We use the well-known theorem in the dynamics about the ellipsoid of inertia [4]. According to this theorem will draw the straight lines l through the center of masses of a line k . We determine in respect of each straight line l the moment of inertia J of the line k and put on

straight lines l from the point O_c vectors with modules

$$|O_c A| = 1/\sqrt{J}. \quad (6)$$

Then the ends of these vectors will lie on the surface of the ellipsoid (see Fig. 1) having a center point O_c .

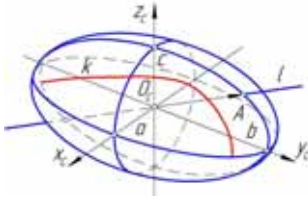


Figure 1: The ellipsoid of inertia of the line k

The ellipsoid of inertia for the center of mass is called the central ellipsoid of inertia. This second order surface characterizes the distribution of the moments of inertia of generator with respect to different axes passing through the center of mass of the generator. The axis of the ellipsoid of inertia are called the principal axes of inertia. These axes are mutually perpendicular. The latter circumstance set the basis for determining the location trihedrons for a projected directing line in the basic coordinate system.

The equation of the ellipsoid of inertia of the line in general form (4) is determined by the components of the inertia tensor. The inertia tensor (3) is the matrix of the quadratic form of the equation of a surface of the second order composed of axial moments of inertia J_{x_c} , J_{y_c} , J_{z_c} and centrifugal moments of inertia $J_{x_c y_c}$, $J_{x_c z_c}$, $J_{y_c z_c}$ with respect to the Cartesian coordinate axes $O_c x_c y_c z_c$ placed at the center of mass of the generator.

In order to extract the geometric information about the value and direction of the axes of the ellipsoid of inertia will make the transition from the equation of the ellipsoid in the general form (4) to the equation in the canonical coordinate system $O_k x_k y_k z_k$ [5]:

$$\frac{x_k^2}{a^2} + \frac{y_k^2}{b^2} + \frac{z_k^2}{c^2} = 1. \quad (7)$$

Take the axis of the central ellipsoids of inertia as the axes of Freinet trihedrons. The unit tangent vector τ is compatible with the smallest of the semi-axes c and headed towards to the next generator. The unit normal vector ν is directed along the semi-axis a . The unit binormal vector β is directed along the semi-axis b . Unit vectors τ , ν , β must be a right-handed vectors. Thus, with the help of the central ellipsoid of inertia is formed the coordinate system (Freinet trihedron) at the center of mass of the curve for a single generator.

We correlated with each of the specified generators the center of mass and Freinet trihedron. Using various types of spline interpolation [6], we can pass from a discrete set of geometric objects obtained bearing the information about projected curve to the continuous spatial curve which is the directing line.

2.2 Forming the spline similar ruled surface strip

The second phase of forming the spline similar ruled surface strip is filling of the surface with continuous carcass of generators defined by the directing line.

Let the equations of the generating segments of discrete carcass in the coordinate system $Oxyz$ have the form

$$x_n = x_n(t), y_n = y_n(t), z_n = z_n(t), \quad (8) \\ 0 \leq t \leq 1,$$

where $n=1, 2, \dots$ – ordinal number of the generating segment, t – parameter.

To construct the fragments will consider each of the n initial generating segments in its own local coordinate system $O_{c_n} x_{c_n} y_{c_n} z_{c_n}$. Local coordinate system we will get by the transferring of origin of the base coordinate system $Oxyz$ in the center of mass O_{c_n} of the corresponding generating segment. In this case axes $O_{c_n} x_{c_n}$, $O_{c_n} y_{c_n}$, $O_{c_n} z_{c_n}$ of the local coordinate systems $O_{c_n} x_{c_n} y_{c_n} z_{c_n}$ are parallel to the axes Ox , Oy , Oz of base coordinate system $Oxyz$. Let us return to the consideration

of the directing line, which is the geometric locus of the vertices of the local coordinate systems. This line will provide direction of forming the fragment of ruled surface strip. Local coordinate systems for the construction of the current generating segments will make the space-parallel movements along the directing line, and the current point of directing line will be the origin of the local coordinate system. Obviously, from point to point will change only the position of the origin, while axes remain parallel to the same name basic axes.

Centers of mass O_{c_n} of the generating segments will be the initial information for designing of the directing line. Thus, by determining the centers of mass of the generating segments, we receive discrete information in the form of coordinate for the projected directing line. This discrete information is converted into a continuous, for example, by using normalized cubic splines [6]. As a result, directing line will be piecewise polynomial of the third degree which provides continuity of the derivatives of the second order at the connection points. Application of this type of spline allows to get the whole contour. It consists of segments of curves described by equations of one class of functions. This contour passes through the centers of mass of the generating segments. Segments of curves of contour keep the equality of tangents and curvature radiuses in the respective centers of mass. In addition to information about the directing line the information allowing to recover the generating segment of the modeled fragment of ruled surface strip in each point of this directing line is necessary. To this end we make the transformation equations generating segments (8) in the base coordinate system $Oxyz$ to the equations obtained in the local systems $O_{c_n}x_{c_n}y_{c_n}z_{c_n}$ corresponding to each n -th generating segment. As a result, we obtain new equations of generating segments $x_{c_n} = x_{c_n}(t), y_{c_n} = y_{c_n}(t), z_{c_n} = z_{c_n}(t),$
 $0 \leq t \leq 1.$ (9)

Using the transformed equations generating

segments, calculate their axial $J_{x_{c_n}}, J_{y_{c_n}}, J_{z_{c_n}}$ and centrifugal $J_{x_{c_n}y_{c_n}}, J_{x_{c_n}z_{c_n}}, J_{y_{c_n}z_{c_n}}$ moments of inertia relative to the associated local coordinate systems $O_{c_n}x_{c_n}y_{c_n}z_{c_n}$. This will allow obtain the equations of the central ellipsoid of inertia of generating segments in a general view (4).

Since the used generators are segments of straight lines then its ellipsoid of inertia degenerate into cylindrical surface of revolution [2]. This explains the transition the above mentioned transition from Frenet trihedrons to spatial-parallel movement of local coordinate systems along the directing line.

Thus to each generating segment corresponds to a certain cylindrical surface of revolution which is its dynamic characteristic (see Fig. 2).

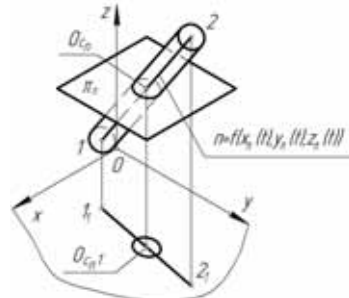


Figure 2: The cylindrical surface of revolution of generating segment

We made the section of each of the cylindrical surfaces of revolution of a plane which parallel to the plane Oxy of the base coordinate system. We obtain the equations of sections representing ellipses

$$J_{x_n}x^2 + J_{y_n}y^2 - 2J_{x_ny_n}xy = 1. \quad (10)$$

Transform equations (10) to the canonical form [5]

$$\frac{x^2}{a_n^2} + \frac{y^2}{b_n^2} = 1 \quad (11)$$

and find the values of large a_n and small b_n

semi-axes of the ellipses (see Fig. 3).

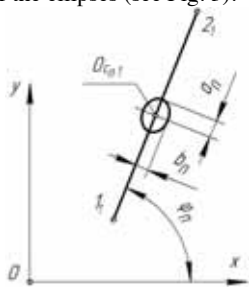


Figure 3: Projections of generating segment and section of the cylindrical surface of revolution on the plane Oxy

We note the following features: 1) the small semi-axis b_n of an ellipse is equal to the radius r_n of a directing circle of the cylindrical surface of revolution in its with a plane; 2) the moment of inertia of the straight line segment relative to the perpendicular axis equals to the value $1/r_n^2$ that follows from the above-mentioned theorem mechanics [4]; 3) moment of inertia of the straight line segment about an axis passing through its center equal to the value $L^3/12$ [2], where L is length of the segment. Of these features we will express relationship between the length L_n of the generating segment and small semi-axis b_n of the ellipse corresponding to this generating segment

$$L_n = 3\sqrt{\frac{12}{b_n^2}}. \quad (12)$$

Using the established relationship (12) we can visually estimate the length of generating segment on one plane projection (see Fig. 3). For the ellipse with a large value of a small semi-axis will correspond a smaller generating segment.

We calculate the angles of slope φ_n of the big semi-axes a_n of ellipses to an axis Ox of base coordinate system through the equation of an ellipse (10) in the general form [5]. With

respect to the spherical coordinates they will azimuth angles.

Zenith angles are the angles of slope the generating segments to the axis Oz . We determined these angles geometrically (see Fig. 3) by using the semi-axes a_n , b_n of the ellipses

$$\varphi_n = \arccos\left(\frac{b_n}{a_n}\right). \quad (14)$$

We transform the discrete information about the azimuth angles φ_n of generating segments and the corresponding axes of the ellipses a_n , b_n in continuous information. We use interpolation by the normalized cubic splines [6]. The parameters of segments of the splines for the directing line and azimuthal angles φ_n and semi-axes of the ellipses a_n , b_n will be the same and change in limits $0 \leq t \leq 1$. Thus the set of parameters $a_m(t)$, $b_m(t)$, $\varphi_m(t)$ will put in correspondence to each point of a directing line, where m – segment number, $m = 1, \dots, n \leq 1$.

We define the length of the generating segments and the values of the corresponding zenith and azimuth angles at each point of the directing line according to the interpolation. This will allow to write the equations of the generator for each segment of a fragment of ruled surface strip in Cartesian coordinates in terms of spherical coordinates

$$\begin{aligned} x_m(t, \varphi) &= \varphi(t) [\varphi(t) [\sin \varphi(t) [\cos \varphi_m(t)]; \\ y_m(t, \varphi) &= \varphi(t) [\varphi(t) [\sin \varphi(t) [\sin \varphi_m(t)]; \\ z_m(t, \varphi) &= \varphi(t) [\varphi(t) [\varphi(t)]. \end{aligned} \quad (14)$$

In equations (14) designations are accepted:

$$\varphi(t) = 3\sqrt{\frac{12}{(b_m(t))^2}}, \quad \varphi(t) = \left(\varphi \frac{1}{2}\right), \quad \varphi(t) = \frac{b_n(t)}{a_n(t)},$$

$$\varphi(t) = \arccos\left(\frac{b_n(t)}{a_n(t)}\right),$$

where $0 \leq t \leq 1$, $0 \leq \varphi \leq 1$.

We will make from the above arguments the equations for segments forming fragment of ruled surface strip in the coordinate-parametric

form in the base coordinate system

$$\begin{aligned}x_m &= \tilde{x}_m(t) + x_m(t, \underline{L}); \\y_m &= \tilde{y}_m(t) + y_m(t, \underline{L}); \\z_m &= \tilde{z}_m(t) + z_m(t, \underline{L}),\end{aligned}\quad (15)$$

where $\tilde{x}_m(t)$, $\tilde{y}_m(t)$, $\tilde{z}_m(t)$ - the current coordinates of the directing line of segment m -th of fragment of ruled surface strip, $x_m(t, \underline{L})$, $y_m(t, \underline{L})$, $z_m(t, \underline{L})$ - the equations of generating segments of m -th segment of fragment of ruled surface strip, $0 \leq t \leq 1$, $0 \leq \underline{L} \leq 1$.

As a result of changes of parameters t and τ of the generating segment it will move continuously along the directing line. The changing segment length depend from the values of parameter b_m . Generating segment continuously rotates relative to the center of local coordinate system depending on the values of the parameters a_m , b_m , \underline{L}_m . Thus at movement the generating segment along the directing line a contiguous fragment of ruled surface strip is formed.

2.3 Development of the theory of the proposed method for the formation of non-ruled strips

Our method can be used for forming non-ruled strips defined discrete by arcs of curves, for example arcs of curves of the second order. These arcs define discrete carcass.

We will define the directing line by a discrete carcass on basis of proposed method. We define the parameters of the ellipsoids of inertia of generators of discrete carcass. We will pass from discrete values of the parameters of ellipsoids to constant values along the length directing line using spline interpolation.

Define equations of generating lines to their respective ellipsoids of inertia, obtained by the interpolation. Draw a straight line through the center of the interpolated ellipsoid of inertia. We take a segment O_cA on the line l (see Fig. 1) from the center of the ellipsoid of inertia to its

intersection with the ellipsoid surface. From the formula (6) follows that the value $J = 1/|O_cA|^2$ is the moment of inertia. Thus the set of values of the moments of inertia with respect to a corresponding set of the straight lines can find by means of the ellipsoid of inertia. The number m of the equation coefficients of generating line requires to calculate m values of moments of inertia for each interpolated ellipsoid. Further we will make system from m the equations according the calculated moments of inertia. Solving the system of equations composed for each interpolated ellipsoid of inertia we obtain the equation of generator which center of mass is located on the directing line.

Thus each point of the directing lines will correspond to certain values of the parameters of ellipsoids of inertia, therefore, will recover the form and position of forming not-ruled strip.

Example 1. Forming the spline similar ruled surface strip

As an example we will consider formation of a fragment of ruled surface strip. Carcass of a required fragment of ruled surface strip is set by six segments (see Fig. 4).

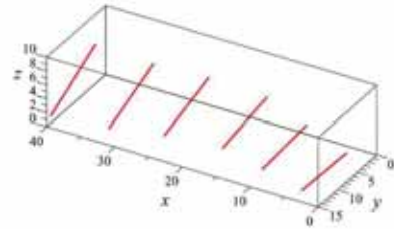


Figure 4: Discrete carcass from the rectilinear generators

Initial data about generate segments are given in Table 1 in the form of mathematical models in the coordinate-parametric form.

Table1: Value of initial data.

Initial data	Value of initial data for forming straight-line segments					
	$n=1$	$n=2$	$n=3$	$n=4$	$n=5$	$n=6$
$x_n(t)$	$1+2\cdot t$	$8+2\cdot t$	$15+2\cdot t$	$23+2\cdot t$	$30+2\cdot t$	$38+2\cdot t$
$y_n(t)$	$8+8\cdot t$	$4+8\cdot t$	$2+8\cdot t$	$3+8\cdot t$	$5+8\cdot t$	$6+8\cdot t$
$z_n(t)$	$3\cdot 2\cdot t$	$4\cdot 3\cdot t$	$5\cdot 4\cdot t$	$6\cdot 5\cdot t$	$7\cdot 6\cdot t$	$8\cdot 7\cdot t$

As the result of calculations by the proposed algorithm the directing line is obtained. It is presented in Figure 5.

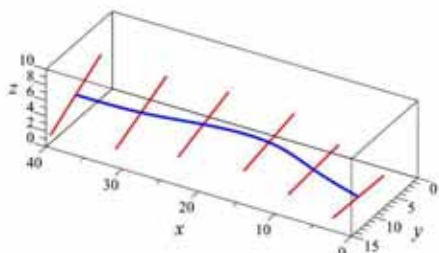


Figure 5: Discrete carcass from the straight-line generators and directing line

As a result forming the fragment of ruled surface strip is obtained. It is presented in Figure 6.

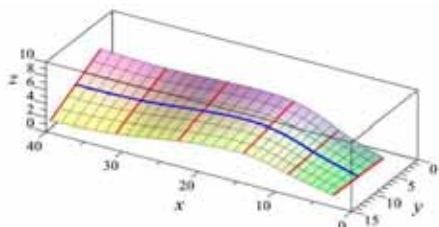


Figure 6: The spline similar ruled surface strip

Example 2. Forming the spline similar parabolic strip

As an example we will consider formation of a fragment of parabolic strip. Let the parametric equations of parabolas in general form in the plane Oxy : $x = p \cdot t^2$, $y = 2[p \cdot t]$, where p – focal parameter of the parabola. Carcass of a required parabolic stripe is set by arcs of three

parabolas in the base coordinate system. This parabolas described by the equations:

$$x_1 = \frac{2}{5}t^2, y_1 = 2[\frac{2}{5}t], z_1 = 0, \quad 0 \leq t \leq 1;$$

$$x_2 = \frac{3}{5}t^2, y_2 = 2[\frac{3}{5}t], z_2 = 0, \quad 0 \leq t \leq 1;$$

$$x_3 = \frac{4}{5}t^2, y_3 = 2[\frac{4}{5}t], z_3 = 0, \quad 0 \leq t \leq 1.$$

The position of the arcs of parabolas in space is defined by shift matrices T_1 , T_2 , T_3 and rotation matrices R_1 , R_2 , R_3 :

$$T_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 2 & 5 & 0 & 1 \end{bmatrix}, \quad T_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 4 & 3 & 8 & 1 \end{bmatrix},$$

$$T_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 6 & 1 & 15 & 1 \end{bmatrix},$$

$$R_1 = \begin{bmatrix} \cos\left(\frac{\pi}{5}\right) & \sin\left(\frac{\pi}{5}\right) & 0 & 0 \\ -\sin\left(\frac{\pi}{5}\right) & \cos\left(\frac{\pi}{5}\right) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$$R_2 = \begin{bmatrix} \cos\left(\frac{\pi}{4}\right) & \sin\left(\frac{\pi}{4}\right) & 0 & 0 \\ -\sin\left(\frac{\pi}{4}\right) & \cos\left(\frac{\pi}{4}\right) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$$R_3 = \begin{bmatrix} \cos\left(\frac{\varphi}{3}\right) & \sin\left(\frac{\varphi}{3}\right) & 0 & 0 \\ -\sin\left(\frac{\varphi}{3}\right) & \cos\left(\frac{\varphi}{3}\right) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

As a result of calculations the equations of m -th segment of spline similar parabolic strip can be written in general form in the basic coordinate system for $m=1, 2$:

$$x_m = \tilde{x}_m(t) + (p_m \varphi^2 - x_{cm}) \cos(\varphi_m) + (2p_m \varphi - y_{cm}) \sin(\varphi_m);$$

$$y_m = \tilde{y}_m(t) + (p_m \varphi^2 - x_{cm}) \sin(\varphi_m) + (2p_m \varphi - y_{cm}) \cos(\varphi_m);$$

$$\tilde{z}_m = \tilde{z}_m(t), \quad 0 \leq t \leq 1, \quad 1 \leq m \leq 1,$$

where $\tilde{x}_m(t)$, $\tilde{y}_m(t)$, $\tilde{z}_m(t)$ - equations of a directing line of a m -th segment of spline similar parabolic strip,

$p_m = f_1(J_{a_m}, J_{b_m}, J_{c_m})$ - focal parameter of a arcs of parabola of a m -th segment,

$x_{cm} = f_2(J_{a_m}, J_{b_m}, J_{c_m})$,

$y_{cm} = f_3(J_{a_m}, J_{b_m}, J_{c_m})$ - coordinates of the center of mass of a arcs of parabola of a m -th,

$\varphi_m = f_4(J_{a_m}, J_{b_m}, J_{c_m}, J_{x_m}(t))$ - angle of rotation of a arcs of parabola of a m -th segment about an axis Oz ,

$$J_{a_m} = 1/(a_m(t))^2, J_{b_m} = 1/(b_m(t))^2,$$

$$J_{c_m} = 1/(c_m(t))^2$$

- axial moments of inertia of a arcs of parabola in the canonical coordinate system about the axes Ox , Oy , Oz , respectively,

$J_{x_m}(t)$ - axial moment of inertia of a arcs of parabola in the basic coordinate system relative to the axis Ox . $J_{x_m}(t)$ was obtained by spline interpolation,

$a_m(t)$, $b_m(t)$, $c_m(t)$ - values of the semi-axes of the central ellipsoid of inertia received as a result of spline interpolation.

In Figure 7 the carcass from arcs of parabolas and a directing line of this carcass is presented.

las and a directing line of this carcass is presented.

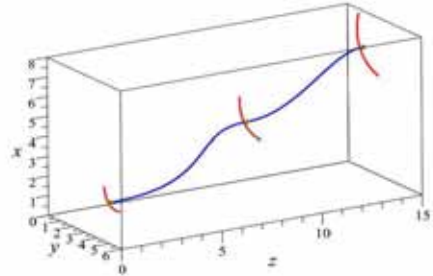


Figure 7: Discrete carcass from arcs of parabolas and directing line

In Figure 8 the formed spline similar parabolic strip is presented.

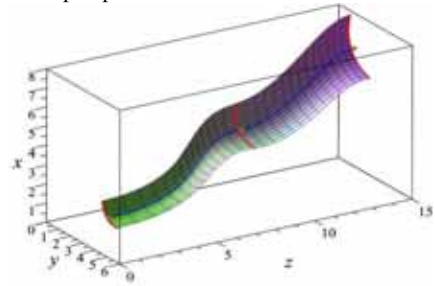


Figure 8: The spline similar parabolic strip

3. CONCLUSIONS

Method of forming spline similar ruled surface strip is proposed in this paper. Method reveals the relationship between geometry and dynamics of the lines. This relationship is established through the moments of inertia, centers of mass and central ellipsoids of inertia of the lines. The possibility of application of the offered method for forming spline similar non-ruled surface strip is shown. Method can be applied in engineering products for forming the transition ruled surfaces if a discrete set of generate segments is given. It can be applied for forming the trough surfaces if a discrete set of arcs of curves of second order also is given.

REFERENCES

- [1] H. Pottmann and J. Wallner. Computational Line Geometry. Berlin, Springer Verlag, Heidelberg, 2001. – 565 p.
- [2] V.V. Dobronravov and N.N. Nikitin. Course of Theoretical Mechanics. Moscow, 1983. – 575 p.
- [3] V.A. Osipov. Machine design methods continuously-frame surfaces. Moscow, 1979. – 248 p.
- [4] N.E. Zhukovsky. Theoretical Mechanics. Moscow-Leningrad, 1952. - 812 p.
- [5] B.N. Delone and D.A. Raikov. Analytic geometry: in 2 volumes. Moscow- Leningrad, 1949. - Volume 1 – 456 p.
- [6] D. Rogers, J. Adams. Mathematical Elements for Computer Graphics. Moscow, 2001. – 604 p.

ABOUT THE AUTHORS

- 1. Denis Sergeevich KORCHAGIN, department of Engineering Geometry & CAD, Omsk State Technical University, Russia.
email: Korch-Den@yandex.ru
- 2. Konstantin Leonidovich PANCHUK, department of Engineering Geometry & CAD, Omsk State Technical University, Russia.
email: Panchuk_KL@mail.ru

FREEFORM SURFACES ADAPTATION USING PLANAR QUADRILATERAL FACETS

Francisco GONZALEZ-QUINTIAL, Antonio SANCHEZ-PARANDIET, and Javier BARRALLO

The University of Basque Country UPV-EHU, Spain

ABSTRACT: Is formulated in this paper the basis of a method of approach to the construction of double-curvature surfaces, through adaptation by the use of developable surfaces.

What arises is a method of adaptation of any forms of double curvature, within the scope of the shapes of the architectural language, to the undoubted technical and above all economic problems posed by the implementation of this type of free form surfaces. It is a process aiming at the simplification of manufacturing systems and hence of the economic impact of these.

The extensive use of this type of free form surfaces of double curvature in the architecture is no doubt during the recent past years, being the search for the construction of this type of shapes with flat or developable elements subject to extensive investigation through multiple ways of approach.

Keywords: Freeform surfaces; Double curved surfaces; Developable surfaces; Algorithmic approach.

1. INTRODUCTION

Architecture has seen in recent years an extensive use of double curvature surfaces. Simultaneously supported in design systems and digital fabrication has been an important development of construction of this type of surfaces that have allowed to see examples of construction of buildings with spectacular results, as Frank Gehry designs supported by the research of J. Glymph [1] geometric developments and D. Shelden [2]. This has led to a revolution within the means of control of less conventional shapes, free form and double curvature shapes not only from the point of view of the CAD technology, but through the impulse that has occurred in the theoretical geometry knowledge and interesting theoretical developments as those published by H. Pottmann [3] intended for specific use in the development of the that it is almost a new science known as Architectural Geometry.

In the works mentioned, a way of leading to resolution of free form shapes approach is its adaptation through developable surfaces, in the case of Shelden what called paper surfaces and in the Pottmann, the generation of planar quad meshes from families of conjugate curves of principal curvature. The reason for the use of this type of planar or developable elements is simplification and economy that offer in regards to constructive development. A developable surface can be constructed from a planar surface, and conversely, a developable surface can be deployed on a plane without any crease or distortion. This question is decisive at the time of manufacture any element forming part of a constructive system either a front facade panel or the hull of a ship. From the geometric point of view there are many aspects that define a developable surface. One of them is its curvature. While a developable surface is characterized by zero Gaussian curvature, on a free surface can be any value equal to or different

from zero Gaussian curvature positive or negative. The adaptation of a freeform surface passes through both by a process of geometric adaptation, where the degree of restrictions posed a developable surface is superior to a freeform surface. We can say that a developable surface is undoubtedly "less free", being subject to important geometrical impositions.

We have developed the basis of a method of adaptation of double curvature surfaces following the proposed philosophy, trying to respond to both sets of adaptation, both through developable strips as using planar facets.

The fundamental difference with the aforementioned methods lies in the former are based mainly on systems built using differential geometry, while what is proposed here follows a path that relies mainly on projective geometry. In addition, as it will be seen, the use of a type of curves very specific, such as the apparent contour curves, as a basis for the design of developable surfaces, let get a result which confers a very particular image, which contains a special link with the perception and the point of view from which are drawn contours, i.e. from which the forms are perceived.

2. GEOMETRIC FOUNDATIONS

When we draw a surface, using any system of representation, whether a hand drawing drawn on paper or a rendering computer model, the process you need to define a curve that mark the edge of the represented object, separating the part of the object to the observer is able to see, which is hidden, or from another point of view, the illuminated area that remains in shadow. This is basically the apparent contour curve or separatrix of that surface. A projecting surface wraps the object that is displayed, depending on the point of view, this surface that is tangent to the object along a series of points that define its apparent contour, will be a developable surface formed by straight lines or projection rays, starting from the point of view made tangent to the surface. These straight lines of projection are the generators of a cone or cylinder.

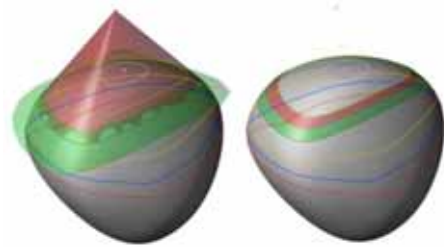


Figure 1, Contours obtained using proper straight trajectory. Projecting conical surfaces.

The theoretical concept is very simple, however the obtaining of these apparent contours is difficult to do especially when it comes to models that move away from simple primitive shapes. Even on primitive forms, obtaining boundary lines of shadow or separatrices, common application in projective geometry of the apparent contours, is laborious when it rushes manually. Computer-aided design programs facilitate and improve this work by providing powerful tools in this regard.

Taking any surface, we can generate a family of apparent contours, through the displacement of the point of view along any spatial trajectory. In our case we will use straight paths either proper Figure 1 or improper ones Figure 2. This process is materialized by obtaining the apparent contour of the surface from a series of points belonging well to a straight line in space, improper straight path case or along a circle of infinite radius, to obtain a improper straight trajectory.

These apparent contours are generated as intersection of projecting surfaces tangent to the base surface, that as noted are absolutely developable surfaces, cones or cylinders, depending on the straight direction of proper or improper projection.

In this way we wrap the primitive surface with a series of tangent developable surfaces which are supported on the surface along the apparent contour and intersect each other and two by two along an outer curve to the surface.

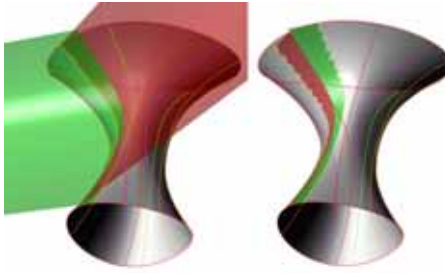


Figure 2. Contours obtained from improper straight trajectory. Projecting cylindrical surfaces.

Projecting tangential surface is bounded by two external curves, intersections between adjacent surfaces and contains the tangent curve on the own projecting surface and the surface base, i.e. the apparent contour. Along each of the apparent contours the generatrix of the projecting surface, cone or cylinder, straight line parallel to the direction of projection slides, generating such developable surface.

In this way we obtain a series of strips that wrap the original surface externally.

These strips are developable, since removed from the projecting tangent surface that defines the apparent contour. Its boundaries are determined by two tangent developable hung abutting surfaces and the breadth of the strip in turn by the distance that separates points of projection from which the tangent surfaces are obtained and with them the apparent contours.

The extension of this document does not allow an exhaustive explanation of geometric principles. This process and the details of the geometric basis are profusely detailed in previous works. [4]

3. GRAPHIC AND PHYSICAL CONSTRUCTION OF DEVELOPABLE STRIPS SUCH ADAPTATION OF DOUBLE CURVATURE SURFACES

The purpose of the method involves obtaining of developable strips that fit surfaces where the curvature is not uniform in regards to their sign, on the other hand this is the usual case in the

complex surfaces or freeform are presented. The success of the strategy focuses on the achievement of a path of adequate projection resulting in satisfactory results in terms of the set of curves of retrieved apparent contour which facilitates the construction of developable strips.

The purpose of the method consists of obtaining developable strips that fit surfaces where the curvature is not uniform in that this is one of the strong points of this system. The path of apparent contours is not subject to any quality of the surface. Depends on fundamentally the situation of the point of view, there are infinite number of apparent contours, since an object can be observed from an infinite number of positions. The observer, the point of view, is what determines its position. The method incorporates this flexibility allowing on the other hand, to be related to a mechanism that is subjective, as it is the vision, geometric resolution of the surface has much to do with appearance or how to look at.

The process will take close proximity, empirically, aided by a system of tracing apparent contours of great agility that allows us to build a generative algorithm in Grasshopper. We can check this way multiple options on the same shape, multi-track, choosing in its case the most appropriate, i.e. the path that family contours optima and within the chosen family select the most favourable contours.

The points of view, in the same way that paths can be chosen arbitrarily, by means of a specific subdivision of the path in a discrete set of points or conversely, can obtain paths passing through certain specific points chosen on the surface, looking for the intersection of the tangent plane to the surface at that point and the line that is being used as path. In this way we can choose different specific contours which pass through certain points and which define singularities in surface or could have any interest for any designing reason. We could also cut off a section of the surface by a given plane containing the path and divide the section produced on the surface. Plotting the tangents at

the various points of division of the intersection we get in the path the vertex of the cone that generate the contour passing through those points.

In the image (Figure 3) we can see the development of the application of the method for the resolution of a surface where the variation of the curvature is noticeable. In this case, used as projection an improper straight trajectory, resulting in apparent contours from tangent surfaces of cylindrical projection. A set of outlines to be selected from points at the infinite, in this case making a regular division of the path, here a circle at infinity, tangent projecting cylindrical surfaces that generate them are obtained. The extraction of the patch bounded by the intersections between cylinders allows to obtain cylindrical developable strip that fits the surface.

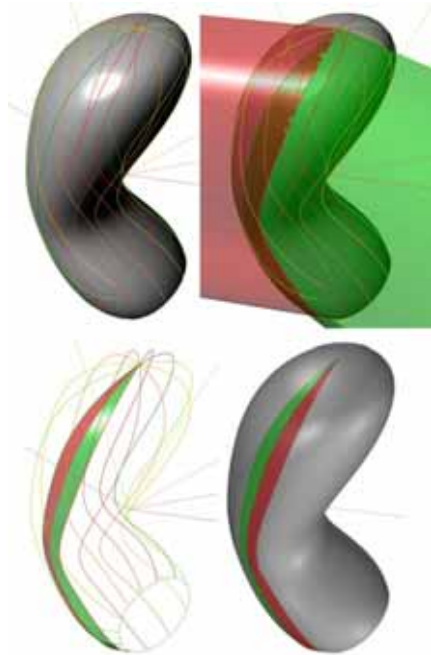


Figure 3. Development of free surface. Application of the method on a free surface by a straight improper trajectory.

In this case because of the nature of the surface, there is the appearance of a node, this is defined as a point on the surface where all paths are cut off. This node is determined by the parallel plane which contains the circular path and is tangent to the surface. Where the surface is closed in the opposite side would appear another node which would determine the straight line which would constitute the axis of rotation of the path with the former. The tangent plane to the surface that will determine this second node must be parallel to the first.

Resolved the graphical mode of adapting the surface of double curvature, it has proceeded to check the validity of the method by transferring to physical models of geometric resolution results. Several cases have been studied from the point of view both graphic and physical, through the construction of prototypes in different materials. We talk about prototypes, not models. On the prototype and during their implementation process, we can anticipate the behaviour of materials in relation to the way in which these are ready, their physical interaction. The prototype has an added-value in relation to purely formal model. We can check the results in terms of tolerances, deformation and adaptability in lightweight materials that make up the prototype and somehow anticipate the behaviour of materials in the final model.

Certainly since they are developable surfaces, it is possible to proceed to the construction of these prototypes by resorting to the use of light laminar materials, it is possible to transfer flat strips development directly using a cutting machine.

As noted, a key feature of the developable surfaces from which we obtain the most useful application in construction, is that in theory, it can be deployed on a plane. It can "develop" without creases or folds, on a flat surface. Conversely, a developable surface can be constructed from a plane, from a flat or planar element we can obtain a developable surface.

It has been able to appreciate, as fidelity achieved in the physical reproduction of the model can be seen in the images Figure 4.



Figure 4. Physical model making applying the process.

As noted initially, the method aims its application to actual scale in the construction of architectural elements. The basic objective of the construction of prototypes has been physically check the reality of the system, the approximation of the shape and the behaviour of the material, as an intermediate step between the graphic and physical process.

4. CONCLUSION

Adaptation is a procedure of alteration of the original shape, so it is necessary to find a balance between constructive economy and fidelity in the reproduction of the form. Based on

these premises has built a system of adaptation of double curvature surfaces using developable strip surfaces, that pretends to be from extensive application to free form surfaces construction. Hereof is the formulation of a geometrical method that as shown it offers satisfactory results, resulting suitable for raised objective, graphic adaptation of form and construction, both prototypes as of end items, that meet the two formulated premises.

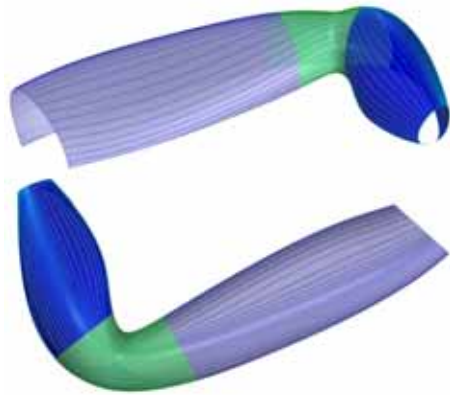


Figure 5. Finding Continuity of contours over freeform surface using different paths

There are important aspects that complement the development and that by extension of this article do not allow to make an exhibition of them. First, it is noteworthy to mention the strategy developed to give continuity to the contours on surfaces which are not continuous, i.e. cutting surfaces and tangents, where the contours, due to the own principles of geometry, lose its continuity. The flexibility of the layout of contours, allows us to combine contours splicing separatrix curves obtained from different points of view. Making a path chaining contour curves, using their intersections along the secant curve to obtain the tangent planes over the adjoining surface. Each of the full

contours obtained on the whole of the area are not produced from the same point of view, but all points of view are located in a same straight trajectory, improper in this case. The result is a family of contours with a remarkable order and continuity over the different surfaces. Figure 5

We referred to the appropriateness of the use of straight paths, there is an important reason linked to the use of this means. The bundle of planes that aims to trace the straight path from which the contours are obtained, cuts off the surface resulting in a series of curves, in this case planar curves, that constitute families of curves that are conjugated to the of apparent contours family ones, that except on quadrics surfaces usually are warped.

Obtaining a network of conjugate curves allows to get on the surface a network whose vertices are on the own surface. This network gives rise to an approximation to the surface in question using quadrilateral planar facets.

This fact allows us to generate a mesh covering the surface by flat elements allowing to use it in approximation or adaptation of the same and therefore its construction, not only using materials with possibility to be curved in one direction from the developable strips, but also using flat rigid materials which do not allow its curvature or bending them represents an economically costly process as it is the case of the glass on the other hand profusely employee and whose process of bending and rolling raises its price in an important way. Figure 6

On the other hand, the fact that the families of conjugated curves to the apparent contours are planar, allows to use them as geometric guidelines of possible structural frames that support the skin of the built form, thus resolving a conflict that we see often in the construction of free-form surfaces where the structure on one side and the covering on the other, usually follow their own geometrical logic.

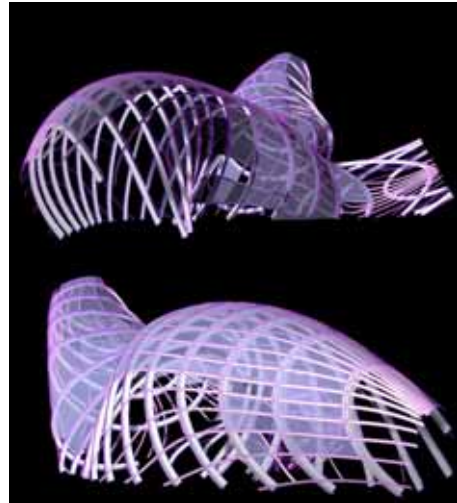


Figure 6. Use of families of conjugate curves for obtaining structural guidelines and flat facets.

REFERENCES

- [1] Glymph, James. Shelden, Dennis. Ceccato, Cristiano. Mussela, Judith. Schoberb, Hans. A Parametric Strategy for Free-Form Glass Structures using Quadrilateral Planar Facets. Proceedings of the 2002 Annual Conference of the Association for Computer Aided Design In Architecture /Pomona (California) 2002
- [2] Shelden, Dennis. Digital Surface Representation and the Constructability of Gehry's Architecture. Thesis Ph.D. Massachusetts Institute of Technology. Dept. of Architecture.
- [3] Pottmann, Helmut. Asperl, A. Hofer, M. Kilian, A. Architectural Geometry. Bentley I. Press. Exton, 2007
- [4] González-Quintial, F. Sanchez-Parandiet, A. Barrallo, J.: 2013, freeform surfaces adaptation through developable surfaces using apparent contours", Proceeding CAAD Futures 2013, Shanghai, China, pp. 358-367

ABOUT THE AUTHORS

1. PhD. Francisco González-Quintial is Architect, Lecturer and Researcher at the School of Architecture in Donostia-San Sebastian (Spain) Building Construction Department.
2. PhD. Antonio Sanchez-Parandiet is Architect, Lecturer and Researcher at the School of Architecture in Donostia-San Sebastian (Spain) Geometry and Graphic Expression Departament.
3. PhD. Javier Barrallo is Lecturer and Researcher at the School of Architecture in Donostia-San Sebastian (Spain) Applied Mathematics Departament. Specialist in Art and Mathematics and Fractal Geometry .

THE GENERAL APPROACH TO THE SOLUTION OF TYPICAL ENGINEERING GEOMETRICAL PROBLEMS

Aleksandr Yurievich BRAILOV

Odessa Academy of Civil Engineering and Architecture, Ukraine

ABSTRACT: In the present work the analysis of the basic methods for solving typical engineering geometrical problems is made. The basic problem is revealed and the first-level tasks of its solution are defined. A general algorithm for solving typical geometrical problems in the form of standardized logical blocks is offered. The general algorithm for the solution of typical geometrical problems is discussed, taking into account the iterative nature of the proposed methods.

Keywords: Method, algorithm, geometry, problem, solution.

1. INTRODUCTION

1.1 Problem statement.

Spatial imagination and imaginative perception are essential professional skills being developed in the academic course "Engineering graphics". The basic sections of engineering graphics are positional problems, metric problems, problems of construction of the development of a curvilinear surface, and problems of constructing of axonometric views of a product [1-13]. These four sorts of problems of the course are called as *typical engineering geometrical problems*.

A number of various methods are used for the solution of typical engineering geometrical problems. As a result, comprehending these methods within the limited course time available is virtually impossible.

Hence, there is a *contradiction* (conflict) between the great variety of known methods for solving typical engineering geometrical problems and the limited course duration. Such a conflict can be partially resolved through the development of a unifying general approach to the solution of typical engineering geometrical problems.

Thus, the problem can be solved by adding offered changes into the existing methodology

of training and engineering practice of the solution of real problems.

The proposed changes will help to solve the following methodological problems:

1. To offer methods and means for increasing the efficiency of comprehension of the theoretical foundations of engineering graphics and the solution of practical problems.
2. To train students in the skills of independent work.
3. To offer graduate students some methods of solving engineering problems related to their specialization in a particular engineering field.

In the conditions of competition, constant increase of requirements to the quality of a product leads to the increase of complexity of real engineering problems under consideration. In such situation known methods of the solution of typical engineering geometrical problems are not always applicable and yield necessary competitive result.

Hence, there is a *contradiction* (conflict) between constant increase of complexity of real engineering problems under consideration and absence of adequate methods of their solution.

The solution of this contradiction (*problem*) can also be partially achieved by working out of the general approach to the solution of typical engineering geometrical problems.

The developed general approach to the

solution of typical geometrical problems in certain extent will allow an engineer to generate algorithm of the solution for novel complex real problem.

1.2 Analysis of researches and publications

Various methods for solving typical engineering geometrical problems are described in the works of Mihajlenko V. E., Vanin V. V., Volkov V. Ya., Kovalev S. N., Najdysh V. M., Podkorytov A. N., Skidan I. A., Ryzhov N. N., Stachel H., Weiss G., Suzuki K. and other scientists [14–27].

Frolov S. A. and Bubennikov A. V. described an algorithm for the solution of positional problems in intersections of geometrical images [17, 14]. This algorithm includes three steps. They also developed another algorithm for the solution of metric problems, consisting of five steps [17, 14]. In their works, three methods of construction of the development of curvilinear and many-faced surfaces were proposed: the normal section method, the development of surface on projection plane method, and the method of triangles [17, 14]. However, algorithm of construction of development for generalizing these methods is not offered. The description of various methods of construction of standard axonometric projections of a product is not presented in the guidelines of the general algorithm.

The knowledge of algorithms facilitates better understanding and comprehension of the essence of methods of the solution of engineering problems correlated with the definition of positional and metric characteristics, constructions of development and axonometric views of geometrical images.

The problem of synthesis of the general algorithm for the solution of all groups of typical engineering geometrical problems was not the object of the current study.

1.3 Aim and purposes of the paper

The purpose of the present paper is to

develop a general formal algorithm for solving typical engineering graphics geometrical problems.

Such an algorithm should help students to understand in depth the geometrical essence of the studied phenomena, to comprehend the generality of engineering methods and to solve practical problems related to the course of study and those in real life more efficiently. Moreover, *the generalized approach* to the solution of typical geometrical problems should also help practical engineers to solve *new* difficult real-life problems.

The paper also aims:

1. To develop a generalized algorithm for solving typical engineering graphics geometrical problems in the form of standard logic blocks, facilitating students' understanding of the essence of the used method for solving problems of descriptive geometry and allowing engineers to solve new difficult real-world problems.

2. To describe the developed general algorithm for the solution of typical geometrical problems, considering the iterative nature of used methods.

2. FOUNDATIONS

The developed general algorithm for the solution of typical geometrical problems consists of seven stages, as shown in Figure 1.

The letters P, M, D, A identifying the types of geometrical problems precede the stage number. Their meaning is as follows: P is for **p**ositional problems; M is for **m**etric problems; D is for problems of construction of **d**evelopment; A is for problems of construction of an **a**xonometric projection of a product. The letter G placed before the stage number indicates the description is considered a stage of the **g**eneral algorithm.

The results of the analysis of the basic methods for solving geometrical problems in the introduced code are shown in Table 1. The content of the introduced symbolic description are discussed in the following sections.

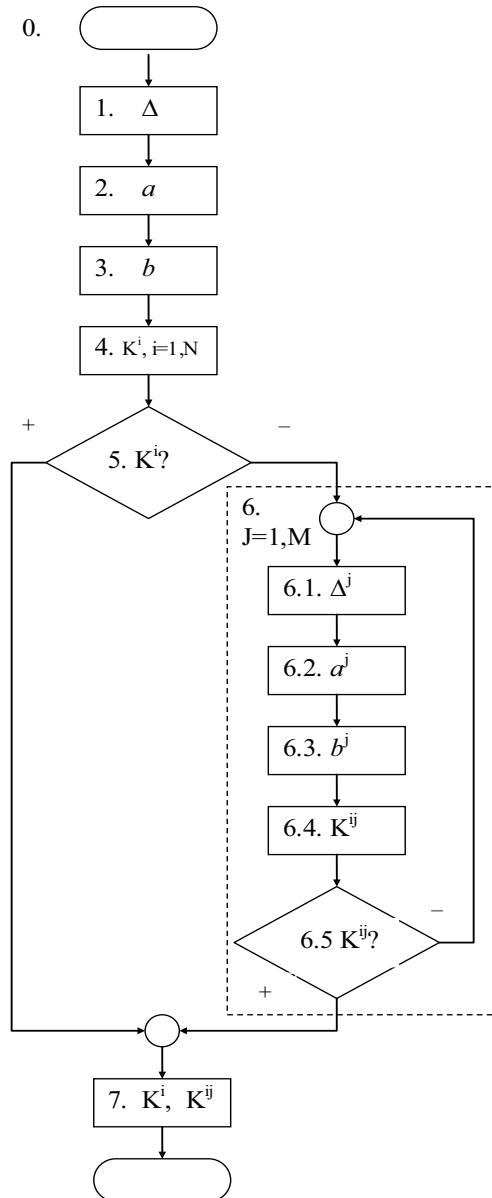


Figure 1: A flow-chart of the algorithm for solving typical geometrical problems.

Table 1: Results of the analysis of the basic methods for solving typical engineering geometrical problems in a symbolic description.

Serial number of a stage of algorithm	Type of an engineering geometrical problem				
	Positional problems	Metric problems	Problems of construction of development	Axonometric problems	General approach New problems
1.	P1	M1	D1	A1	G1
2.	P2	M2	D2	A2	G2
3.	P3	M3	D3	A3	G3
4.	P4	M4	D4	A4	G4
5.	P5	M5	D5	A5	G5
6.	P6	M6	D6	A6	G6
7.	P7	M6	D7	A7	G7

P1. According to the criterion of simplicity of the intersecting lines of the auxiliary geometrical image (intermediary) with the original geometrical images *the kind of this auxiliary geometrical image (intermediary)* is imaginarily figured out.

On the basis of imaginarily-defined three-dimensional representation of the auxiliary geometrical image, *construction of its two-dimensional image* is carried out. The result of the first stage is the complex drawing of the intermediary — $\Delta(\Delta_1, \Delta_2)$.

M1. The number of the chosen systems of planes of projections for allocation and a designation of the end result of the solution of a metric problem is defined as $j=M$. Number M belongs to a set of natural numbers.

At the beginning, to the counter of the chosen systems of planes of projections Δ^j , $j=1, M$ the first value unit is given — $j=1$.

The projection of a geometrical image and the axis of coordinates for removal from initial complex drawing $K^i(K^i_l, K^i_m)$, $i=1, N$, $l=1, 4, 5$, $m=2, 4, 5$ for a new plane of projections Δ^j (Δ^j_l , $l=4, 5$, Δ^j_m , $m=4, 5$, $j=1, M$) are defined.

The coordinate axis X_{lm} , $l \neq m$ is then constructed with respect to the preserved projection of a geometrical image so that the geometrical image occupied its particular position in relation to the introduced plane of projections — Δ^j (Δ^j_l , $l=4, 5, 6$; Δ^j_m , $m=4, 5, 6$;

$l \neq m$).

D1. On a curvilinear developed surface between two nearest generating lines a certain region Δ is allocated.

A1. *The orthogonal projection* Δ (the front view, the top view, the left-side view etc.) containing the maximum information on design features of a product (an aperture, a lug, planes, facets, flutes etc.) is highlighted in the complex projective drawing (or in a prototype photographs).

Three orthogonal projection planes ($j = 0, 1, 2$) are normally used in the axonometry construction.

As such, the counter j considered orthogonal planes of projections Δ^j , $j=0, M$ gets its first value zero $j=0$. Number M belongs to a set of natural numbers.

G1. Thus, *the auxiliary image* Δ is defined at the first stage of the solution of a typical geometrical problem.

P2. *The first auxiliary line* a of intersecting of the intermediary Δ with the first initial geometrical image is constructed $\Sigma — a = \Delta \cap \Sigma$. The result of the second stage is the complex drawing of the first auxiliary line $a — a(a_1, a_2)$.

M2. Lines of projective connections, which remain from the saved projection of the geometrical image, are constructed to be perpendicular to the created axis of coordinates — $a \perp X_{lm}$, $l \neq m$ (a^n , $n=1, 2, 3, 4, \dots$).

D2. The curvilinear surface located between the selected generating lines is replaced with a flat figure — *a*.

A2. A geometrical analysis of the orthogonal projection Δ keeping the maximum quantity of images of the constructive elements is carried out. The aim of this analysis is the definition of the characteristic geometrical images (circles, arches, segments of straight lines, points).

In order to construct an axonometry of the characteristic geometrical images their characteristic points are allocated — $K^i(K^i_1, K^i_2, K^i_3)$, $i=1, N$. Number N belongs to the set of natural numbers. For example, the axonometric projection of a segment of a straight line can be constructed using two characteristic points $K^i(K^i_1, K^i_2, K^i_3)$, $i=1, 2$.

The primary *a* coordinates x^i, y^i, z^i , $i=1, N$ are defined for each characteristic point $K^i(x^i, y^i, z^i)$, $i=1, N$ belonging to the analyzed orthogonal projection Δ .

G2. Thus, at the second stage of the solution of a typical geometrical problem for the auxiliary image Δ , the first group of auxiliary actions *a* is carried out.

P3. The second auxiliary line *b* of intersection of the same intermediary Δ with the second initial geometrical image Φ is constructed — $b=\Delta \cap \Phi$. The result of the third stage for a positional problem is the complex drawing of the second auxiliary line *b* — $b(b_1, b_2)$.

M3. For a metric problem, marks *b* are placed on the constructed lines of projective connections from the created axis of coordinates to a created projection of a geometrical image to visualize the distances accordingly equal to distances from the deleted projection to the deleted axis of coordinates — $b(b^n, n=1, 2, 3, 4, \dots)$.

D3. For a problem of construction of development, the true size of the constructed flat figure is defined — *b*.

A3. At the third stage of construction of an axonometry of a product by means of the

natural coefficients of distortion k_x, k_y, k_z for coordinate axes Ox, Oy, Oz , secondary axonometric coordinates x'_K, y'_K, z'_K of the axonometric projection K' into the primary orthogonal coordinates x_K, y_K, z_K points $K'(x^i, y^i, z^i)$, $i=1, N$ are defined:

$$x'_K = k_x \cdot x_K, \quad y'_K = k_y \cdot y_K, \quad z'_K = k_z \cdot z_K.$$

The result *b* of the third stage for the problem of construction of an axonometry are the secondary axonometric coordinates x'_K, y'_K, z'_K defined for each characteristic point — $(K'(x'_K, y'_K, z'_K))^i$, $i=1, N$.

G3. Thus, at the third stage of the solution of a typical geometrical problem for an auxiliary image Δ , the second group of auxiliary actions *b* is carried out.

P4. Points of intersection K^i , $i=1, N$ of the first auxiliary line *a* and the second auxiliary line *b* are defined — $K^i=a \cap b$, $i=1, N$. Numbers N belong to the set of natural numbers. The results of the fourth stage are a complex drawings of points of intersecting $K^i=a \cap b$, $i=1, N$ the first auxiliary line *a* and the second auxiliary line *b* — $K^i(K^i_1, K^i_2)$, $i=1, N$.

M4. The projection of the geometrical image is constructed by connecting characteristic points K^i , $i=1, N$ and the constructed projection of a geometrical image is designated (indicated) — $K^i(K^i_l, K^i_m)$, $i=1, N$, $l=1, 4, 5$, $m=2, 4, 5$, $l \neq m$.

The discussed algorithm can be used many times for the same engineering metric problem if the following two conditions are met.

Conditions of repeated application of a projections planes change method.

One of the two planes of projections of the replaced system should be a component of the introduced system of planes of projections.

The entered plane of projections should be perpendicular to the saved plane of projection.

D4. The flat figure of the true size is constructed on the plane (drawing) — K^i , $i=1, N$. Number N is the number of identical flat figures of the true size corresponding to the allocated part Δ of a curvilinear developed surface.

A4. Axonometric projections of the

characteristic points are constructed — $(K'(x'_k, y'_k, z'_k))^i, i=1, N$.

Connecting axonometric projections $(K'(x'_k, y'_k, z'_k))^i, i=1, N$ of the characteristic points $K^i(x^i, y^i, z^i), i=1, N$ by a thin line, the constructed projections of the geometrical images belonging to an orthogonal projection Δ of a product are marked.

The result of the fourth stage includes the axonometric projections $(K'(x'_k, y'_k, z'_k))^i, i=1, N$ of the characteristic points $K^i(x^i, y^i, z^i), i=1, N$ and the axonometric projections of the geometrical images belonging to the orthogonal projection Δ of a product.

G4. Thus, at the fourth stage of the solution of a typical geometrical problem for an auxiliary image Δ , the necessary projection of the geometrical image $K^i, i=1, N$ is constructed.

P5. Sufficiency of the obtained number of points of intersecting $K^i=a \cap b, i=1, N$ of the first auxiliary line a and the second auxiliary line b for allocation and designation of the end result of solution of a positional problem is verified.

If the number of points of intersecting $K^i=a \cap b, i=1, N$ of the first auxiliary line a and the second auxiliary line b for obtaining the end result is sufficient, then the last seventh stage of the given algorithm is carried out.

If the number of points of intersecting $K^i=a \cap b, i=1, N$ of the first auxiliary line a and the second auxiliary line b for obtaining the end result is insufficient, then the following sixth stage of the developed algorithm is carried out.

M5. Sufficiency of the obtained number of the new introduced systems of planes of projections for the allocation and designation of the end result of the solution of the problem (Fig. 1) is verified.

In doing so, the new constructed projection using the characteristic points $K^i, i=1, N$ is checked out for the purpose of possible definition of the required metric characteristic of the initial geometrical image.

If the number of the new entered systems of planes of projections is sufficient, then the last seventh stage of the developed algorithm is

carried out. Otherwise, the sixth stage of the developed algorithm is carried out.

D5. The verification of the completion of construction of the development of the surface is carried out.

If all identical flat figures of the true sizes corresponding to the allocated fragment Δ of the curvilinear developed surface coincide with a plane and other fragments $(\Delta^j, j=1, M)$ are not allocated, the seventh stage is carried out. Otherwise, the sixth stage of the developed algorithm is carried out.

A5. Sufficiency of the obtained number of axonometric projections $(K'(x'_k, y'_k, z'_k))^i, i=1, N$ of characteristic points $K^i(x^i, y^i, z^i), i=1, N$ for the allocation and designation of the end result of the solution of an axonometric problem is verified.

If the number of axonometric projections $(K'(x'_k, y'_k, z'_k))^i, i=1, N$ of characteristic points $K^i(x^i, y^i, z^i), i=1, N$ is sufficient, then the last seventh stage of the developed algorithm is carried out. For example, for a flat plate its axonometric projection can be allocated by a thick line. Otherwise, the sixth stage of the developed algorithm is carried out.

G5. Thus, at the fifth stage of the solution of a typical geometrical problem for an auxiliary image Δ sufficiency of the obtained result is verified.

P6. Additional points $K^{ij}, i=1, N, j=1, M$ for unequivocal allocation and a designation of the end result of the solution of a positional problem are defined.

The counter j of the additional intermediaries $\Delta^j, j=1, M$ is given the first unit value $j=1$.

P6.1. The complex drawing of a new intermediary $\Delta^j, j=1, M$ is constructed — $\Delta^j(\Delta^j_1, \Delta^j_2), j=1, M$.

For this new intermediary $\Delta^j, j=1, M$ steps P6.2, P6.3, P6.4 of the developed algorithm are carried out in full analogy with steps P2, P3, P4.

P6.2. The complex drawing of a new auxiliary line $a^j, j=1, M$ for the new intermediary $\Delta^j, j=1, M$ is constructed — $a^j(a^j_1, a^j_2)$.

P6.3. The *complex drawing* of a *new auxiliary line* b^j , $j=1, M$ for the new intermediary Δ^j , $j=1, M$ is constructed — $b^j(b^j_1, b^j_2)$.

P6.4. *Complex drawings of new points of intersection* $K^{ij}=a^j \cap b^j$, $i=1, N$, $j=1, M$ for auxiliary lines a^j , b^j , $j=1, M$ are constructed — $K^{ij}(K^{ij}_1, K^{ij}_2)$, $i=1, N$, $j=1, M$.

P6.5. A *decision* on the *unambiguity* of the construction of the end *result* of the solution of a problem is *made*.

If the number of points K^{ij} , $i=1, N$, $j=1, M$, belonging simultaneously to two initial geometrical images and various intermediaries Δ^j , $j=1, M$ is sufficient for the unequivocal allocation and designation of the result of the solution of a positional problem, the seventh stage of the developed algorithm is carried out.

Otherwise, the value of the counter j additional intermediaries Δ^j , $j=1, M$ is increased by unit ($j=j+1$). A complex drawing of a new intermediary Δ^j , $j=1, M$ is constructed and new points of intersection $K^{ij}=a^j \cap b^j$, $i=1, N$, $j=1, M$ of new auxiliary lines a^j , b^j , $j=1, M$ are defined. In other words, the sixth stage of the developed algorithm is repeated.

M6. *Additional points* K^{ij} , $i=1, N$, $j=1, M$ of the new geometrical image in the new system of planes of projections for unequivocal allocation and designation of the end result of the solution of a metric problem are *defined*.

The counter j new entered planes of projections Δ^j , $j=1, M$ is assigned with a new value as $j=j+1$.

M6.1. The *complex drawing* of a *new system* of planes of projections for a new entered plane of projections Δ^j , $j=1, M$ is constructed — $\Delta^j(\Delta^j_1, \Delta^j_m)$, $j=1, M$.

The projection of a geometrical image and the axis of coordinates deleted from the initial complex drawing for the introduced plane of projection are defined — $\Delta^j(\Delta^j_1, \Delta^j_m)$, $j=1, M$.

The coordinate axis X^j_{lm} , $l \neq m$, $j=1, M$ is constructed relative to the kept projection of the geometrical image so that the geometrical image occupies a particular position in relation

to the introduced plane of projections — $\Delta^j(\Delta^j_1, \Delta^j_m)$, $j=1, M$.

For the introduced plane of projections Δ^j , $j=1, M$ steps M6.2, M6.3, M6.4 the given algorithm are carried out in full analogy with steps M2, M3, M4 points of the developed algorithm.

M6.2. *New lines* a^j , $j=1, M$ of projective connections from the saved projection of the geometrical image are *constructed* to be perpendicular to the created axis of coordinates for the newly introduced plane of projections Δ^j , $j=1, M$ — $a^j(a^n_1, a^n_m, n=1, 2, \dots)$; $a^j \perp X^j_{lm}$, $l \neq m$, $j=1, M$.

M6.3. The *distances* accordingly equal to distances from the deleted projection to the deleted axis of coordinates b^j , $j=1, M$ for the new entered plane of projections Δ^j , $j=1, M$ are *laid down* by notches on the constructed lines of projective connections from the introduced axis of coordinates to the constructed projection of the geometrical image — $b^j(b^{jn}_1, b^{jn}_m, n=1, 2, \dots)$.

M6.4. A *newly constructed projection* of the geometrical image of the new entered plane of projections Δ^j , $j=1, M$ is *allocated* through connecting the characteristic points K^{ij} , $i=1, N$, $j=1, M$ and *designated* — $K^{ij}(K^{ij}_1, K^{ij}_m)$, $i=1, N$, $j=1, M$.

M6.5. A *decision* about the *unambiguity* of the construction of an end *result* of the solution of the problem is *made*.

If the number of points K^{ij} , $i=1, N$, $j=1, M$ of newly constructed projection of the geometrical image for the newly introduced plane of projections Δ^j , $j=1, M$ is sufficient for unequivocal allocation and designation of the metric result of the solution of the problem, then the seventh stage of the developed algorithm is carried out.

Otherwise, the value of the counter j of the newly entered plane of projections Δ^j , $j=1, M$ increases by unit ($j=j+1$).

The complex drawing of a new plane of projections Δ^j , $j=1, M$ is constructed and new points K^{ij} , $i=1, N$, $j=1, M$ for the newly obtained

plane of projections are defined for a new geometrical image, that is, the sixth stage of the given algorithm is repeated.

D6. Steps D6.1, D6.2, D6.3, D6.4, D6.5 of the developed algorithm are *executed* as steps D1, D2, D3, D4, D5 for the next part of the curvilinear surface — Δ^j , $j=1, M$.

To the counter j of this next part Δ^j , $j=1, M$ of the curvilinear surface the first unit value $j=1$ is assigned.

D6.1. The *next part* of the curvilinear surface is *allocated* — Δ^j , $j=1, M$. This allocated part directly borders with the previously allocated part of the curvilinear surface.

D6.2. The curvilinear surface located between the allocated generating lines is *replaced* with a *flat figure* — a^j , $j=1, M$.

D6.3. *True size* of the obtained flat figure is *determined* — b^j , $j=1, M$.

D6.4. The *flat figure* of the true size is *constructed* on a plane (drawing) — K^{ij} , $i=1, N$, $j=1, M$.

D6.5. *Verification* of the termination of the construction of the development of the surface is *carried out* — K^{ij} , $i=1, N$, $j=1, M$.

If all identical flat figures of the true size corresponding to the allocated part Δ^j , $j=1, M$ of the curvilinear developed surface coincide with a plane, and other parts (Δ^j , $j=1, M$) are not going to be allocated, then the seventh stage is carried out.

If all identical flat figures of the true size corresponding to the allocated part Δ^j , $j=1, M$ of the curvilinear developed surface coincide with a plane, and other parts (Δ^j , $j=1, M$) are going to be allocated, then the sixth stage of the developed algorithm is carried out.

The value of the counter j the allocated parts Δ^j , $j=1, M$ is increased by the unit value ($j=j+1$). The complex drawing of an approximating flat figure a^j , $j=1, M$ is constructed. The true size of the obtained flat figure is determined — b^j , $j=1, M$. The flat figure of the true size is constructed on a plane (drawing) — K^{ij} , $i=1, N$, $j=1, M$. That is, the sixth stage of the developed algorithm is repeated.

A6. *Additional points* K^{ij} , $i=1, L$, $j=1, M$ for unequivocal allocation and designation of the end result of the solution of the axonometric problem are *defined*. Number L belongs to a set of natural numbers.

To the counter j , a second unit value is assigned for the next orthogonal projection Δ^j , $j=1, M$ of the product — $j=1$.

A6.1. An *orthogonal projection* is *allocated* in the complex projective drawing (or prototype photos). This orthogonal projection Δ^j , $j=1, M$ (the front view, the top view, the left-side view, etc.) contains the maximum information on the design features of the product (an aperture, a lug, planes, facets, flutes, etc.) in comparison with *remained orthogonal projections* Δ^j , $j=2, M$.

For the *newly allocated orthogonal projection* Δ^j , $j=1, M$, steps A6.2, A6.3, A6.4 the developed algorithm are carried out in full analogy with steps A2, A3, A4.

A6.2. A *geometrical analysis* of a new orthogonal projection Δ^j , $j=1, M$ is *carried out* for the definition of characteristic geometrical images (circles, arches, pieces of straight lines, points).

For the construction of axonometry of the characteristic geometrical images, their *characteristic points* are *allocated* — K^{ij} ($K^{ij}_1, K^{ij}_2, K^{ij}_3$), $i=1, L$; $j=1, M$.

The number L of new characteristic points K^{ij} ($K^{ij}_1, K^{ij}_2, K^{ij}_3$), $i=1, L$; $j=1, M$ does not include *characteristic points* K^i (K^i_1, K^i_2, K^i_3), $i=1, N$ for which *axonometric projections* ($K'(x'_k, y'_k, z'_k)$), $i=1, N$ have been already *constructed*.

The *original coordinates* x^{ij}, y^{ij}, z^{ij} , $i=1, L$; $j=1, M$ are *defined* for each characteristic point K^{ij} (x^{ij}, y^{ij}, z^{ij}), $i=1, L$; $j=1, M$ of the analyzed orthogonal projection Δ .

Thus, on the considered iteration of the solution of a typical geometrical problem for an auxiliary image Δ^j , $j=1, M$, the *first group of auxiliary actions* a^j , $j=1, M$ is *carried out*. As a result of these actions a^j , $j=1, M$, the *original coordinates* x^{ij}, y^{ij}, z^{ij} , $i=1, L$; $j=1, M$ of the

additional characteristic points K^{ij} ($K^{ij}_1, K^{ij}_2, K^{ij}_3$), $i=1,L; j=1,M$ of the product are defined.

A6.3. At this stage of construction of axonometry of the product by means of natural coefficients of distortion k_x, k_y, k_z for coordinate axes Ox, Oy, Oz , the secondary axonometric coordinates $(x'_k, y'_k, z'_k)^{ij}$, $i=1,L; j=1,M$ of the axonometric projection $(K')^{ij}$, $i=1,L; j=1,M$ into primary orthogonal coordinates $(x_k, y_k, z_k)^{ij}$, $i=1,L; j=1,M$ points $K^{ij}(x^{ij}, y^{ij}, z^{ij})$, $i=1,L; j=1,M$ are defined as:

$$x'_k = k_x \cdot x_k, \quad y'_k = k_y \cdot y_k, \quad z'_k = k_z \cdot z_k.$$

The results b^j , $j=1,M$ of this stage for the problem of construction of axonometry are the secondary axonometric coordinates $(x'_k, y'_k, z'_k)^{ij}$, $i=1,L; j=1,M$ for each characteristic point $K^{ij}(x^{ij}, y^{ij}, z^{ij})$, $i=1,L; j=1,M$ — $(K'(x'_k, y'_k, z'_k))^{ij}$, $i=1,L; j=1,M$.

Thus, at this stage of the solution of a typical geometrical problem for the auxiliary image Δ^j , $j=1,M$, the second group of auxiliary actions b^j , $j=1,M$ is carried out. The result of these actions b^j , $j=1,M$ is the determination of the secondary axonometric coordinates $(x'_k, y'_k, z'_k)^{ij}$, $i=1,L; j=1,M$ of characteristic points $K^{ij}(x^{ij}, y^{ij}, z^{ij})$, $i=1,L; j=1,M$ of the product.

A6.4. Axonometric projections of the characteristic points are constructed — $K^{ij}(x^{ij}, y^{ij}, z^{ij})$, $i=1,L; j=1,M$.

Connecting the axonometric projections $(K'(x'_k, y'_k, z'_k))^{ij}$, $i=1,L; j=1,M$ of the characteristic points $K^{ij}(x^{ij}, y^{ij}, z^{ij})$, $i=1,L; j=1,M$ by a thin line, one can allocate the constructed projections of the geometrical images belonging to the orthogonal projection Δ^j , $j=1,M$ of the product.

A result of the fourth stage includes the axonometric projections $(K'(x'_k, y'_k, z'_k))^{ij}$, $i=1,L; j=1,M$ of the additional characteristic points $K^{ij}(x^{ij}, y^{ij}, z^{ij})$, $i=1,L; j=1,M$ and allocated by a thin line the additional axonometric projections of the geometrical images belonging to an orthogonal projection Δ^j , $j=1,M$ of the product.

Thus, the fourth stage of the solution of a

typical geometrical problem for the auxiliary image Δ^j , $j=1,M$ includes the construction of the additional necessary projections of geometrical images.

A6.5. A decision on the unambiguity of the construction of the end result for the problem is made.

If quantities of points K^{ij} , $i=1,N+L; j=1,M$, belonging to geometrical images of orthogonal projections Δ^j , $j=0,M$ is enough for unequivocal allocation and a designation of result of the solution of an axonometric problem, the seventh stage of the given algorithm is carried out.

Otherwise, the value of the counter j of the additional orthogonal projections Δ^j , $j=2,M$ is increased by the value unit ($j=j+1$). The new orthogonal projection Δ^j , $j=2,M$ is allocated. The auxiliary procedures a^j , b^j , $j=1,M$ are carried out. After this, a decision on the unambiguity of construction of the end result is made. That is, the sixth stage of the developed algorithm is repeated.

G6. Thus, at the sixth stage of solving a typical geometrical problem additional points K^{ij} , $i=1,N; j=1,M$ of the new constructed geometrical image for unequivocal allocation and a designation of the end result for solving the problem are defined.

P7. The result of the solution of the positional problem is allocated and designated by the connection of the obtained points K^{ij} , $i=1,N; j=1,M$ of threefold incidence by a smooth line taking into account their visibility on the complex drawing.

M7. The result of the solution of the problem, i.e. a geometrical image with the demanded metric characteristic, is allocated and designated. The size of the required metric characteristic is determined.

D7. All the constructed that far adjacent flat figures are allocated with smooth curves and designated — K^i , $i=1,N; K^{ij}$, $i=1,N; j=1,M$. The constructed set of adjacent flat figures is an approximate development of the considered curvilinear surface.

A7. The result of the solution of the

axonometric problem is *allocated* and *designated* by connection of obtained points K^{ij} , $i=1, N+L+P$, $j=1, M$ by a smooth line taking into account their visibility on the projective.

The connecting and overall dimensions are added to the constructed axonometry of the product.

G7. Thus, at the seventh stage of the solution of a typical geometrical problem *the end result* is *allocated* and *designated*.

P. The developed flow-chart of the algorithm of solution of positional problems on mutual intersecting of geometrical images reflects repetition of the first five stages at the sixth stage for new *intermediaries* Δ^j , $j=1, M$ [2].

M. The developed flow-chart of the algorithm of solution of metric problems reflects repetition of first five stages at the sixth stage for new *planes of projections* Δ^j , $j=1, M$ [4].

D. The developed flow-chart of the algorithm of the construction of a development of a surface reflects repetition of first five stages at the sixth stage for the next *fragment of a curvilinear surface* — Δ^j , $j=1, M$ [3].

A. Thus, the developed flow-chart of the algorithm of the construction of axonometry of the product reflects repetition of first five stages at the sixth stage for new *orthogonal projections* Δ^j , $j=1, M$ (Fig. 1).

G. Thus, the developed flow-chart of the general algorithm of the solution of typical geometrical problems reflects repetition of first five stages at the sixth stage for new *images* Δ^j , $j=1, M$ (Fig. 1).

The flow-chart of the general algorithm of the solution of typical geometrical problems (Fig. 1) corresponds to the flow-chart of the algorithm of the solution of positional problems on mutual intersecting of geometrical images [2], the flow-chart of algorithm of the solution of metric problems on definition of the demanded metric characteristic [4], the flow-chart of algorithm of construction of a development of a surface [3] and the flow-chart of the algorithm of the construction of an axonometry of a product.

The symbolical designations used in blocks are kept identical to facilitate the complete comparison of the flow-chart of the general algorithm of the solution of typical geometrical problems (Fig. 1), the flow-chart of the algorithm of the solution of positional problems on mutual intersecting of geometrical images [2], the flow-chart of the algorithm of the solution of metric problems [4], the flow-chart of the algorithm of construction of development of a surface [3] and the flow-chart of the algorithm of construction of an axonometry of the product.

The developed algorithm flow-chart contains only standard logic blocks of operators for programming using a language of high level.

Efficiency of the general algorithm is verified by many students and engineers dealing with the solution of practical problems.

For example, using the algorithm of the proposed structure, students solve the positional problem of the construction of the intersecting line of two surfaces of rotation (Fig. 2).

Results of the solution of the problem are shown in a Fig. 3.

In tutorial study and/or in an exam, students with the help of the developed algorithm of structure can solve in a timely manner a metric problem of determining of the true size of segment ΔABC (Fig. 4) of a plane.

Results of solution of a problem are shown in Figure 5 for three-dimensional and two-dimensional geometrical models.

For example, using the algorithm of the developed structure in practical training of students, the problem of construction of the development of a curvilinear surface of inclined circular cone S (Fig. 6) can be solved.

Students receive result for taken away time.

Results of solving of a problem are shown together with approximation of geometrical model of inclined circular cone S by tetrahedral model of a pyramid (Fig. 7).

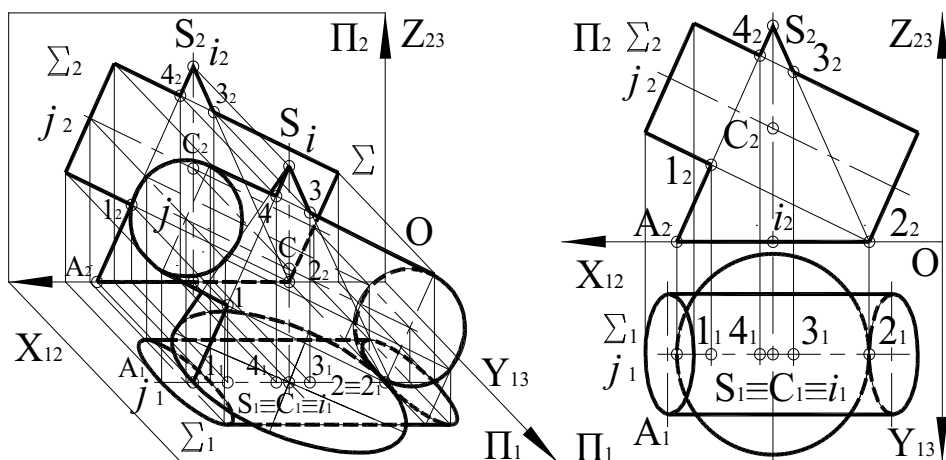


Figure 2: Three-dimensional and two-dimensional models of initial geometrical images of a cone and a cylinder.

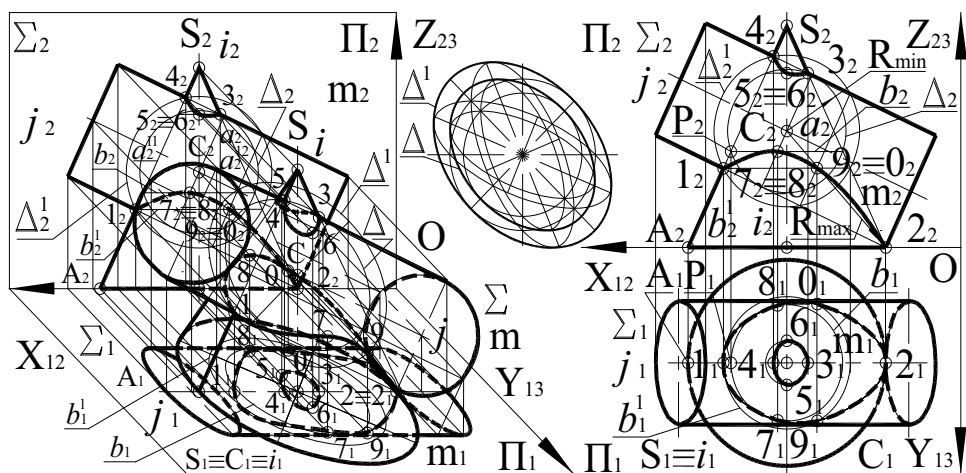


Figure 3: Geometrical models of the construction of the line of intersection of two surfaces of rotation.

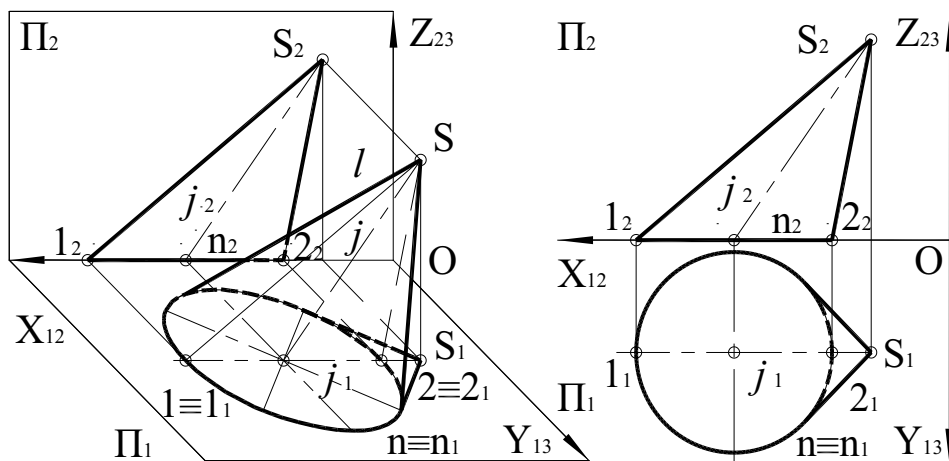


Figure 6: Geometrical models of inclined cone S.

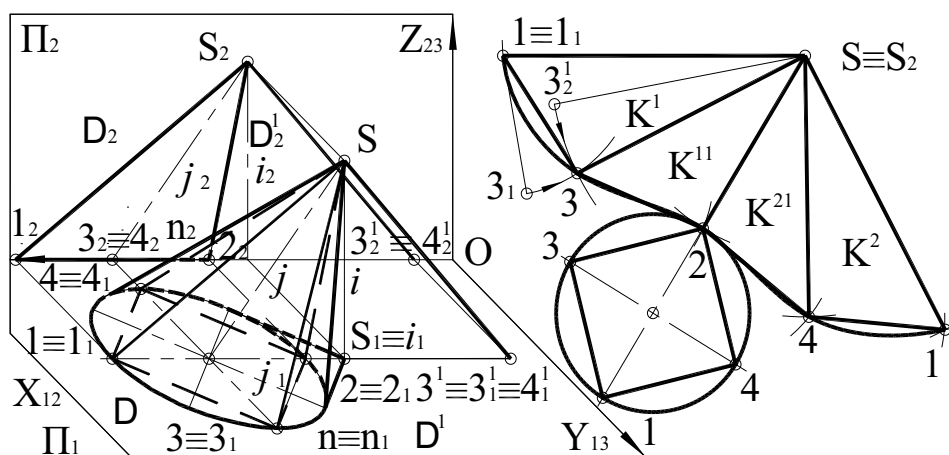


Figure 7: The development of a curvilinear surface of an inclined circular cone.

3. CONCLUSIONS

3.1. The developed algorithm *generalizes* known approaches of solving typical geometrical problems [5].

3.2. The flow-chart of general algorithm for solving typical geometrical problems (Fig. 1)

corresponds to a flow-chart of algorithm for solving positional problems on mutual intersecting of geometrical images [2], a flow-chart of algorithm for solving metric problems by defining the necessary metric characteristic [4], a flow-chart of algorithm for constructing development of a surface [3], and

a flow-chart of algorithm for constructing the axonometric views of a product [13].

3.3. Using proposed general algorithm *in the form of standard logic blocks* facilitates students' and specialists' understanding of a geometrical essence of the studied phenomenon and an essence of the method for solving a problem.

3.4. The developed general approach to the solution of typical geometrical problems considers *the iterative nature* of used methods.

3.5. The introduced formal general algorithm allows students to master their *skills* of independent work.

3.6. Practical engineers can use the developed general approach for solving *new* difficult real-world problems.

3.7. *Synthesized general algorithm*, as the *main contribution* to geometry, allows, on *deductive basis* – "from the general to quotient", to teach and study Engineering Geometry.

REFERENCES

- [1] Brailov A. Yu. Features of training on engineering graphics in modern conditions // *Technical esthetics and design* (in Russian). – K: Vipol, 2011. – Issue 8. – P. 44–49.
- [2] Brailov A. Yu. The structure of algorithm of the solution of positional problems // *Applied geometry and the engineering graphics* (in Russian). – K.: KNUBA, 2011. – Issue 88. – P. 100–105.
- [3] Brailov A. Yu. The structure of algorithm of the construction of development of a surface // *Applied geometry and the engineering graphics* (in Russian). – K.: KNUBA, 2012. – Issue 89. – P. 94–100.
- [4] Brailov A. Yu. The structure of algorithm of the solution of metric problems // *Works of Tavrijsky state agrotechnological university* (in Russian). – Melitopol: TSATU, 2013. – SPGM-15. – P. 16–24.
- [5] Brailov A. Yu. The general algorithm of the solution of typical geometrical problems // *Applied geometry and the engineering graphics* (in Russian). – K.: KNUBA, 2013. – Issue 91. – P. 32–45.
- [6] Brailov A. Yu. "Interference in design", (in Russian) *Proc. 10th ICGG*, Kiev, Ukraine, 2002, Vol. 1, 84–88.
- [7] Brailov A. Yu. "Design using T-FLEX CAD", *Proc. 11th ICGG*, Guangzhou, China, 2004, 397–402.
- [8] Brailov A. Yu. "On the development of a parametrical three-dimensional model of a product", *Proc. 12th ICGG*, Salvador, Brazil, 2006, Paper #A19.
- [9] Brailov A. Yu. *Computer engineering graphics in the environment of T-FLEX: Transformations of two-dimensional and three-dimensional models of products* (in Russian). – Kiev: Caravella, 2007. – 176 p.
- [10] Brailov A. Yu. A theoretical approach to transformations of two-dimensional and three-dimensional models of the product // *Proceedings of the Thirteenth International Conference on Geometry and Graphics* (Dresden, GERMANY). – ISGG, 2008. – P. 58–59.
- [11] Brailov A. Yu. Fundamental principles of the design and technological development of an engineering product // *Proceedings of the Fourteenth International Conference on Geometry and Graphics* (Kyoto, JAPAN). – ISGG, 2010. – P. 324–325.
- [12] Brailov A. Yu. Laws of projective connections // *Proceedings of the Fifteenth International Conference on Geometry and Graphics* (Montreal, CANADA). – ISGG, 2012. – P. 16–17.
- [13] Brailov A. Yu. *Engineering Geometry* (in Russian). – Kiev: Caravella, 2013. – 456 p.
- [14] Bubennikov A. V. *Descriptive geometry* (in Russian) / A. V. Bubennikov – M.: Vishaya shkola, 1985. – 288 p.
- [15] Cardone V., Iannizzaro V., Barba S., Messina B. Computer aided descriptive geometry // *Proceedings of the Fifteenth*

- International Conference on Geometry and Graphics*, Montreal, CANADA. – ISGG, 2012. – P. 100—109.
- [16] Honma I. A trial with teaching materials on descriptive geometry by using CAD for students with hearing impairments // *Proceedings of the Fifteenth International Conference on Geometry and Graphics*, Montreal, CANADA. – ISGG, 2012. – P. 296—301.
- [17] Frolov S. A. *Descriptive geometry* (in Russian), Moscow, Mashinostroenie, 1978. – 240 p.
- [18] Mihajlenko V. E. *Engineering and computer graphics* (in Russian). / V. E. Mihajlenko, V. M. Najdich, A. N. Podkorytov, I. A. Skidan – Kiyv: Vishcha shkola, 2001. – 350 p.
- [19] Mihajlenko V. E., Vanin V. V., Kovalyev S. N. *Engineering and computer graphics* (in Russian). – K: Karavella, 2013. – 328 p.
- [20] Podkorytov A. N., Galzman E. G., Perevalov V. F. *Lectures on engineering graphics (with structurally logic schemes and algorithms of graphic constructions in solving typical problems) for students of non-mechanical specialties* (in Russian). – Odessa: OSPU, 1993. – 83 p.
- [21] Ohtsuki M., Ohtsuki A. Descriptive geometry and graphical user interface // *Proceedings of the Fifteenth International Conference on Geometry and Graphics*, Montreal, CANADA. – ISGG, 2012. – P. 563—568.
- [22] Ryan D.L. *CAD/CAE descriptive geometry* / Daniel L. Ryan. – Boca Raton: CRC Press, 1992. – 209 p.
- [23] Schmitt F. Descriptive geometry and 3D-CAD // *Proceedings of the Eleventh International Conference on Geometry and Graphics* (Guangzhou, CHINA). – ISGG, 2004. – P. 257—262.
- [24] Stachel H. Descriptive geometry meets computer vision – the geometry of multiple images // *Proceedings of the Twelfth International Conference on Geometry and Graphics* (Salvador, BRAZIL). – ISGG, 2006. – Paper #T30.
- [25] Suzuki K., Fukano A., Kanai T., Kashiwabara K., Kato, M., Nagashima S., Tanaka I., Tsutsumi E., Yokoyama Yu., Adachi H., Kondo K., Yamaguchi Ya. Development of graphics literacy education (2) – Full implementation at the university of Tokyo in 2007 // *Proceedings of the Thirteenth International Conference on Geometry and Graphics* (Dresden, GERMANY). – ISGG, 2008. – P. 228.
- [26] Volkov V. Ya., Yurkov V. Yu., Panchuk K., Ilyasova O., Kaygorodtseva N., Yakovenko K. The innovative paradigm of teaching in descriptive geometry // *Proceedings of the Fifteenth International Conference on Geometry and Graphics*, Montreal, CANADA. – ISGG, 2012. – P. 778—787.
- [27] Weiss G. Is advanced elementary geometry on the way to regain scientific terrain? // *Proceedings of the Fifteenth International Conference on Geometry and Graphics*, Montreal, CANADA. – ISGG, 2012. – P. 93—804.

ABOUT THE AUTHORS

Aleksandr Yurievich Brailov, PhD, Dr. Sci., Professor, is a Professor of the Department of Descriptive Geometry, Engineering Graphics and Computer Aided Design; Odessa Academy of Civil Engineering and Architecture. His research interests are Mechanical Design, Computer Aided Design, Engineering Computer Graphics and Computer Design of Tools. He can be reached by telephone number: +38-(048) 766-59-39, +38-(095) 864-11-81; by e-mail: brailov@gmail.com, mailing address: Akademika Filatova, 4, Korpus, A, Kvartira, 30, Odessa-80, 65080, UKRAINE.

GENERALIZED METALLIC MEANS FAMILY

Vera Martha Winitzky de SPINADEL
The University of Buenos Aires, Argentina

ABSTRACT: The Metallic Means Family, introduced by the author in 1998, is the set of positive solutions of the quadratic equation $x^2 - px - q = 0$, where p and q are natural numbers.

The infinite members of the MMF are classified in two subfamilies:

- 1) PPMF which are the positive solutions of equation $x^2 - px - 1 = 0$ that have a purely periodic continued fraction expansion $x = [n, n, n, \dots]$. The most important member is the Golden Mean $\phi = 1,618\dots$
- 2) PMMF which are the positive solutions of equation $x^2 - x - q = 0$ that have a periodic continued fraction expansion (not purely periodic).

Applying George Odom's method of finding the Golden Ratio in an equilateral triangle and a cube, we have extended the search of analogue ratios occurring at n -simplices and polytopes. Analyzing the sequences of ratios, we found members of the MMF and new ones for p and q real numbers. We propose to call this larger set of Metallic Means, the Generalized Metallic Means Family

Keywords: Golden Mean, n -simplices, polytopes, Metallic Means family.

1. INTRODUCTION

The Metallic Means Family (MMF) was introduced by the author in 1998 [1], [2]. The most prominent member of this family is the well known Golden Mean. Among its relatives, let us mention the Silver Mean, the Bronze Mean, the Copper Mean, the Nickel Mean, etc. The MMF is the set of positive quadratic irrational numbers σ_p^q which are solutions of the equation

$$x^2 - px - q = 0 \quad (1)$$

where p and q natural numbers. It is divided into two subfamilies:

- a) the positive solutions σ_p^1 of equation (1) with $q = 1$

$$x^2 - px - 1 = 0 \quad (2)$$
- b) the positive solutions σ_1^q of equation (1) with $p = 1$

$$x^2 - x - q = 0. \quad (3)$$

It is easy to verify that all the members of σ_p^1 have a purely periodic continued fraction expansion of the form

$$x = [n, n, \dots] = \overline{[n]} \quad (4)$$

In fact, if we take $p = 1$, we obtain $x^2 - x - 1 = 0$, which can be written $x^2 = x + 1$ and dividing by $x \neq 0$, the result is the equation $x = 1 + \frac{1}{x}$. Replacing iteratively

the value of x we obtain

$$x = \overline{[1]} = \sigma_1^1 \quad (5)$$

that is a purely periodic continued fraction

which is equal to $\phi = \frac{1 + \sqrt{5}}{2} = 0,618\dots$, the well known Golden Mean.

Taking $p = 2$ we find the value

$$\sigma_2^1 = \overline{[2]} \quad (6)$$

called the Silver Mean $\sigma_{Ag} = 1 + \sqrt{2} = 2,414....$

For $p = 3$ we have

$$\sigma_3^1 = \left[\bar{3} \right] = \frac{3 + \sqrt{13}}{2} = 3,3027... \quad (7)$$

known as the Bronze Mean σ_{Br} and so on. This subfamily is called the PPMF (purely periodic Metallic Means Family) because the continued fraction expansions of all are purely periodic.

With respect to the second subfamily, which members are denoted by σ_1^q , it is easy to verify that they have a periodic continued fraction expansion, not purely periodic, of the form

$$\sigma_1^q = \left[m, \overline{n_1, n_2, \dots} \right]. \quad (8)$$

Some of the members of this subfamily, denoted by PMMF (periodic Metallic Means Family) are integers. The first one $\sigma_1^2 = \left[2, \bar{0} \right] = 2$ is called the Copper Mean and the next integers appear in quite a regular way (See Reference [2]).

The first non-integer Metallic Mean is obtained solving the quadratic equation $x^2 - x - 3 = 0$, which positive solution is

$$x = \frac{1 + \sqrt{13}}{2} = \left[2, 3, \bar{\dots} \right] = \left[2, \bar{3} \right]$$

a periodic continued fraction expansion that is called the Nickel Mean..

All the non-integer Metallic Means of the PPMF share the property that the continued fraction expansions of the non-integer Metallic Means are “palindromic” about their centers, except for the last digit of the period, which equals $2m - 1$.

All the members of the MMF enjoy common mathematical properties that are fundamental in the actual research in every type of design, on the stability of micro- and macro-physical systems, going from the DNA geometrical internal structure to the astronomical galaxies, on the search of universal roads to chaos in the analysis of non linear dynamical systems and so on.

2. APPLICATION OF ODOM'S METHOD

Let us begin trying to find the “Golden Mean” in any geometric figure or even in spaces of dimension greater than 2. Instead of starting with a golden rectangle as usual, we could adopt the method introduced by George Phillips Odom Jr. He was born in 1941 and gathered some recognition early in his career for his light machines made from fiber optics that he exhibited at the *Knoll International Gallery* in Manhattan, USA, in the 1960s. Later, his career faltered somewhat and he could not repeat his early success. He suffered from depressions which culminated in a suicide attempt and a subsequent hospitalization at the Hudson River Psychiatric Center in Poughkeepsie, USA, where he became a permanent residence since the early 1980s.

Odom found out that the Golden Mean occurs in an equilateral triangle and its circumcircle. Upon hearing it from Odom, the well known Canadian geometer H. S. M. Coxeter submitted it as a problem to the *American Mathematical Monthly* (Problem E2007, 1983), [3].

The result of Odom was obtained connecting the midpoints A and B of two sides of an equilateral triangle and intersecting this line with the circum-circle, obtaining points

X and Y , see Fig.1a. Then the ratio $\frac{AB}{BX} = \phi$, as it is easy to prove.

In effect, if the triangle PQR has sides of length $2a$ then we have

$$\overline{PB} = \overline{BR} = a;$$

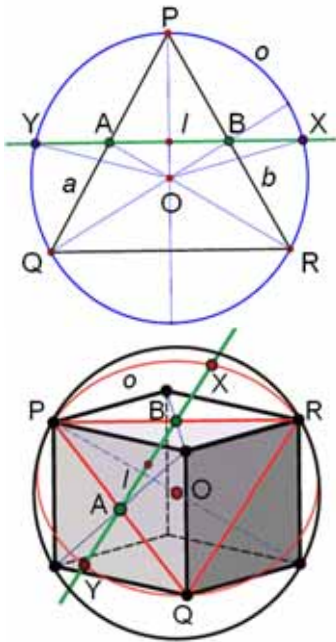
$$\overline{YA} = \overline{BX} = b.$$

Applying the theorem of chord intersections we see that

$$\overline{BY} \cdot \overline{BX} = \overline{PB} \cdot \overline{BR} \rightarrow$$

$$(a+b)b = a.a$$

If $a/b = x$ then $1 + x = x^2$ and $x^2 - x - 1 = 0$ which positive solution is the Golden Mean.



Figures 1a) and 1b)

Odom also recognized that the intersection of the line connecting the midpoints of two adjacent square faces of a cube with the circum-sphere of the cube leads to the Golden Ratio too as can be seen at Fig. 1b.

$$\frac{\overline{AX}}{\overline{AB}} = \frac{\overline{AB}}{\overline{BX}} = \phi$$

3. GENERALIZATION TO N-SIMPLICES

One of the possibilities to generalize G. Odom's method could be to consider n-simplices. An n-simplex is the smallest convex set that contains the given vertices. For example:

0-simplex: point

1-simplex: line segment

2-simplex: triangle

3-simplex: tetrahedron

The method would consist in connect the centers A, B of adjacent faces and intersect this line with the circumhypersphere of the simplex, using a unit hypersphere.

Let us start with a regular tetrahedron, as it is shown in Figure 2. The points V, W are the midpoints of opposite edges of the tetrahedron.

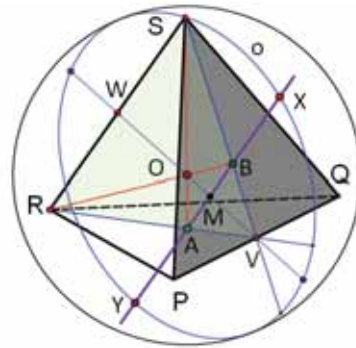


Figure 2. Odom's idea applied to a regular tetrahedron.

$$\begin{aligned}\overline{PQ} &= u := 2\sqrt{\frac{2}{3}}, \quad \overline{OV} = \sqrt{\frac{1}{3}}, \quad \overline{OM} = \frac{1}{3}\overline{OV} \\ \overline{MX}^2 &= 1 - \overline{OM}^2 = 1 - \frac{1}{27} = \frac{26}{27}, \quad \overline{AB} = 2\overline{MA} = \frac{u}{3}, \\ \overline{XA} &= \overline{MX} + \overline{MA}, \quad \overline{MA} = \frac{1}{3}\overline{VA} = \frac{1}{3}\sqrt{\frac{2}{3}}, \quad \overline{YA} = \overline{MX} - \overline{MA},\end{aligned}$$

Now we can calculate the simple ratios $R(XBA)$ and $R(YBA)$; and then the double cross ratio $CR(YXBA) = R(YBA)/R(XBA) = r_{X,Y}$

$$\begin{aligned}\Rightarrow R(XBA) &= \frac{\overline{XA}}{\overline{BA}} =: r_x = \frac{1}{2}(1 + \sqrt{13}), \quad R(YBA) = \frac{\overline{YA}}{\overline{BA}} =: r_y = \frac{1}{2}(-1 + \sqrt{13}), \quad CR(YXBA) = 3, \\ \Rightarrow r_{x,y} &\dots \text{solution of } x^2 - x - 3 = 0.\end{aligned}$$

This gives an idea of how to calculate the double cross ratio $r_{X,Y}$ also for n -simplices, using Figure 2 and Figure 3 as a guideline.

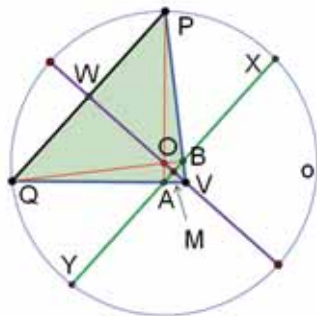


Figure 3: Cross section of an n -simplex through an edge PQ .

The altitude \overline{PA} of an n -simplex has length $\frac{1}{n}\sqrt{3}$ and is divided by the centre O with ratio $R(O, A, P) = -1:n$; that of the $(n-1)$ -faces has length $\frac{1}{n-1}\sqrt{3}$ and is divided by its center A with ratio $-1:(n-1)$. Therewith it follows that the simple ratio ρ_x and the analogue ratio ρ_y are the solutions of the quadratic equation $x^2 - x - \frac{n(n-1)}{2} = 0$.

For $n = 2, 3, 4, 5, \dots$ the ratios $\rho_x(n)$ are equal to: $\phi, \frac{1}{2}(1 + \sqrt{13}), 3, \frac{1}{2}(1 + \sqrt{41}), \dots$ and all the quadratic equations $x^2 - x - q(n) = 0$ have integer coefficients $q(n) = 1, 3, 6, 10, \dots$

4. GENERALIZATION TO POLYTOPES

A polytope is a polyhedron of dimension greater than 3. For $n = 4$ there exists 6 regular polytopes:

- 1) the simplex, that has 5 vertices, 10 edges, 10 triangular and 5 tetrahedral faces
- 2) the hypercube, with all cubic faces and 4 edges in each vertex.
- 3) the 16-cells or hexadecocoron with 8 vertices, 24 edges, 32 triangular and 16 tetrahedral faces. In each vertex there are octahedrons.
- 4) the 24 cells or icositetrahedron with 24 vertices, 96 edges, 96 triangular and 24 octahedral faces. In each vertex, it has cubes.
- 5) the 120-cells or hecatonicosacoron with 600 vertices, 1.200 edges and 720 pentagonal and 120 dodecahedral faces. In each vertex it has tetrahedrons.
- 6) the 600 cells or hexacosicoron with 120 vertices, 720 edges and 1.200 triangular and 600 tetrahedral faces. In each vertex it has icosahedrons.

For $n > 4$, there are 3 regular polytopes:

- a) the simplex
- b) the hypercube
- c) the cross-polytope

It is interesting to mention that quite recently, at the Bridges Towson 2012 Art exhibition, the mathematician Henry Segerman presented "Dual Half 120- and 600-Cells" 2011.

PA 220 Plastic, Selective-Laser-Sintered, 3 sculptures each 10.5 cm x 10.5 cm x 10.5 cm. (see Reference [4]).

In all the previous cases, we can apply Odom's procedure but it is necessary to make a difference between the case of taking into account the centers A and B of opposite faces or of adjacent faces of the geometric figure. Last year, together with Prof. Gunter Weiss, we have presented a regular conference at the VII Mathematics & Design International Conference, which was held at the Faculty of Architecture and Urbanism of the province of San Miguel de Tucumán, Argentina, from 02 to 06 September 2013.

In this conference, which will be published in a special issue of the Journal of Mathematics and Design, with the Proceedings of it, we have considered the application of Odom's method to hypercubes, octahedrons and cross-polytopes, obtaining in all the cases second degree equations of the type:

$$x^2 - x - q = 0$$

with q a real number, some of which positive solutions are members of the PPMF.

With respect to the other subfamily of the PPMF, it is easy to verify that if we make the transformation $x \rightarrow y$, where

$$y^2 = x^2/q$$

we find the second degree equation

$$y^2 - \frac{1}{\sqrt{q}} y - 1 = 0$$

which is of the type

$$x^2 - px - 1 = 0$$

where p is a real number and among its positive solutions, we will encounter the members of the subfamily PPMF.

5. GMMF

Of course, we are not able to consider the continued fraction expansions of all the members of this GMMF because now we have real values for the coefficients p and q but we may proceed to the generalization considering a Fibonacci generalized sequence. Let us recall that a Fibonacci sequence is a sequence of natural numbers determined by taking each number equal to the sum of the last two. For this reason, this type of sequences are called "*secondary Fibonacci sequences*", to distinguish them from the ternary Fibonacci sequences, in which each term is a linear combination of the last three.

Beginning with $F(0) = 1$; $F(1) = 1$, we have the following sequence

$$1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, \dots \quad (9)$$

where

$$F(n+2) = F(n+1) + F(n).$$

As it is well known

$$\lim_{n \rightarrow \infty} \frac{F(n+2)}{F(n+1)} = \phi \quad (10)$$

The Fibonacci sequence may be generalized, giving birth to "*generalized secondary Fibonacci sequences*" (GSFS) of the type

$$a, b, pb + qa, p(pb + qa) + qb, \dots$$

which satisfy relations

$$G(n+1) = p G(n) + q G(n-1)$$

were p and q are real numbers.

If we calculate the ratio

$$\frac{G(n+1)}{G(n)} = p + q \frac{G(n-1)}{G(n)} = p + \frac{q}{\frac{G(n)}{G(n-1)}}$$

and take limits in both members of this equation, assuming that $\lim_{n \rightarrow \infty} \frac{G(n+1)}{G(n)}$ exists and is equal to a real number x -- fact that will be proved in next theorem--, we have

$$x = p + \frac{q}{x}$$

or

$$x^2 - px - q = 0,$$

which positive solution is

$$x = \frac{p + \sqrt{p^2 + 4q}}{2}.$$

Theorem

Given a generalized secondary Fibonacci sequence (GSFS)

$$a, b, pb + qa, p(pb + qa) + qb, \dots$$

such that

$$G(n+1) = p G(n) + q G(n-1),$$

$p, q \in \mathbb{R}$, then there exists $\lim_{n \rightarrow \infty} \frac{G(n+1)}{G(n)}$ and is equal to a real positive number σ .

Proof:

To find the n th term of the GSFS, let us put

$$\begin{aligned} G(n+1) &= p G(n) + q H(n) \\ H(n+1) &= G(n) \end{aligned}$$

and

$$\overline{G(n)} = \begin{pmatrix} G(n) \\ H(n) \end{pmatrix}; A = \begin{pmatrix} p & q \\ 1 & 0 \end{pmatrix}.$$

Then it is easy to prove that

$$\overline{G(n+1)} = A \cdot \overline{G(n)}.$$

Let us assume that $G(0) = G(1) = 1$ for simplicity.

$$\text{If } \overline{G(1)} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \text{ then } \overline{G(n+1)} = A^n \cdot \overline{G(1)}$$

and the problem is reduced to finding the n th power of the matrix A .

We know that the eigenvalues of A are

$$\sigma = \frac{p + \sqrt{p^2 + 4q}}{2}; \sigma' = \frac{p - \sqrt{p^2 + 4q}}{2}.$$

To diagonalize A so as to transform it in the form $A_d = \begin{pmatrix} \sigma & 0 \\ 0 & \sigma' \end{pmatrix}$, we use the change of

$$\text{base matrix } P = \begin{pmatrix} \sigma & \sigma' \\ 1 & 1 \end{pmatrix}.$$

The n th power of A is calculated applying the similarity transformation

$$A^n = P \cdot A_d^n \cdot P^{-1} = \frac{1}{\sigma - \sigma'} \begin{pmatrix} \sigma^{n+1} - \sigma'^{(n+1)} & \sigma\sigma'(\sigma^n - \sigma'^n) \\ \sigma^n - \sigma'^n & \sigma\sigma'(\sigma^{(n-1)} - \sigma'^{(n-1)}) \end{pmatrix}$$

and the n th term of the GSFS

$$1, 1, p + q, p(p + q) + q, \dots$$

is given by the following expression

$$G(n+1) = \frac{\sigma^{n+2} - \sigma'^{(n+2)}}{\sigma - \sigma'}.$$

Replacing $\sigma - \sigma' = \sqrt{p^2 + 4q}$; $\sigma' = -\frac{q}{\sigma}$
we obtain

$$\lim_{n \rightarrow \infty} \frac{G(n+1)}{G(n)} = \lim_{n \rightarrow \infty} \frac{\sigma^{n+1} + \left(-\frac{q}{\sigma}\right)^{n+1}}{\sigma^n + \left(-\frac{q}{\sigma}\right)^n} = \sigma$$

and the proof is completed.

Note 1: if instead of choosing $G(0) = G(1) = 1$ we begin the GSFS with two arbitrary values a and b , it is easy to prove that the result is the same. Indeed, given the GSFS

$$a, b, pb + qa, p(pb + qa) + qb, \dots$$

we have to evaluate the quotient.

$$\frac{G(n+1)}{G(n)} = \frac{pnG(n) + qaG(n-1)}{pbG(n-1) + qaG(n-2)} = \frac{pb \frac{G(n)}{G(n-1)} + qa}{pb + \frac{qa}{\frac{G(n-1)}{G(n-2)}}} \xrightarrow{n \rightarrow \infty} \sigma$$

which tends to the value of x when $n \rightarrow \infty$.

Note 2: we have already applied this result in the case of the MMF with p and q natural numbers (see Reference [2], p. 195).

6. CONCLUSIONS

Even when we are aware that the number of members of this GMMF is infinite, there are certain differences in their mathematical behavior which are fundamental in many applications. Let us mention the main results of this new research:

- The members of the GMMF intervene in the determination of the quasi-periodical behavior of non linear dynamical systems, being therefore an invaluable key in the search of universal ways on the roads from order to chaos.
- The numerical sequences based on the members of this family satisfy many additive properties and simultaneously are geometric sequences, having the possibility of being taken as the base of many new systems of proportions, appropriate to be used on every type of design.

REFERENCES

- [1] Spinadel, Vera W. de, “La familia de Números Metálicos en Diseño”, 1er. Seminario Nacional de Gráfica Digital. Sesión de Morfología y Matemática. FADU, UBA, Argentina, vol. II, pp. 173-179, ISBN 960-25-0424-9, 11-13 Junio 1997.
 - [2] Spinadel, Vera W. de, “From the Golden Mean to Chaos”. Book edited by Nueva Librería, Argentina. 1998, ISBN 950-43-9329-1. 2nd Edition, Editorial Nobuko in 2004, ISBN 987-1135-48-3. 3rd Edition, Nueva Librería in 2010, ISBN 978-987-1104-83-3.
 - [3] Coxeter, H.S.M.: Problem E2007, *American Mathematical Monthly*, 1983.
 - [4] Carter Paul, REPORT Bridges Towson 2012 Art Exhibition, Journal of Mathematics and the Arts, 2013. Vol. 7, No. 1, pp. 40-49.
- <http://dx.doi.org/10.1080/17513472.2013.765327>
- [5] Weiss Gunter and Spinadel, Vera W. de, “From George Odom to a new system of Metallic Means” Proc. of the VII International Mathematics & Design Conference (02-06 September 2013), Faculty of Architecture and Urbanism FAU, Tucumán, Argentina. Published as a Special Issue of the Journal of Mathematics & Design, 2014.

ABOUT THE AUTHOR

Vera Winitzky de Spinadel obtained a PhD in Mathematics at the Faculty of Exact Sciences, University of Buenos Aires (UBA) in 1958. At present, she is Full Emeritus Professor of Mathematics at the Faculty of Architecture, Design and Urban Planning, UBA, since 2010. In 1995, she was designed Director of the research Centre of Mathematics & Design MAyDI. In April 2005, she inaugurated the Laboratory of Mathematics & Design at the University City, 3rd Pavilion, 4th floor, Buenos Aires. Argentina. Since 1995, she has organized every three years several international conferences of Mathematics & Design. She is the president of the International Mathematics & Design Association since 1998, being one of its objectives to publish a Journal of Mathematics & Design. She is author of fifteen books and has published more than hundred original papers at national and international reviews.

Her e-mail is: vspinade@fibertel.com.ar
vspinade@gmail.com
myd_lab@yahoo.com.ar
<http://www.maydi.org.ar>

GENERATING THE GROWTH MODEL GEOMETRY BASED ON THE HUMAN POSTURE

Shun IWASAWA and Yoichiro KAWAGUCHI
The University of Tokyo, Japan

ABSTRACT: This paper presents a new method for generating geometry of the GROWTH model for an interactive artwork. The GROWTH model was originally presented by Yoichiro Kawaguchi [1]. The rendering methods presented in this paper is based on previous researches [2-4], but with improvements as follows: 1) Capturing the human posture data including the angle of arm joints, tilt of torso and the position of the whole body in the sensing area, 2) Relating such data to the parameters determining the model geometry, and 3) Introducing a new coloring method so-called “branch direction dependent color” for model’s surface property.

Keywords: GROWTH model, Motion Capture, Interactive Artworks, GPU Computing.

1. INTRODUCTION

A method proposed in this paper is intended to be used in the field of “computer graphics interactive arts”. For the purposes of this paper, this term will be taken to mean artworks that capture movements of dancer or audience (hereinafter referred to as “participant”) by using sensing devices and change computer graphics imagery interactively. It is general for such artworks to have data captured by sensors to be input as the parameters for procedural model causing a morphology formation of the graphics. In particular, by introducing the procedural model which is derived from natural phenomenon such as growth, propagation or evolution of the natural life forms, the generated CGI could gain some lifelike characteristics. Therefore, the artwork can expect to give participants an impression of “communicating” with virtual life.

We chose the “GROWTH model” which is considered one of the pioneers in the field of procedural model. “GROWTH model” (Figure 1) is the process for generating the forms of nature based on spirals and ramification by using the algorithmic structure. This model can create not only the existing shapes of nature, but also the variety of imaginary organic shapes

and movements.

Additionally, the models are usually rendered with “Metaballs” (geometric isosurface), which enable the model to take on the biotic flexibility. That is another remarkable feature of this model.

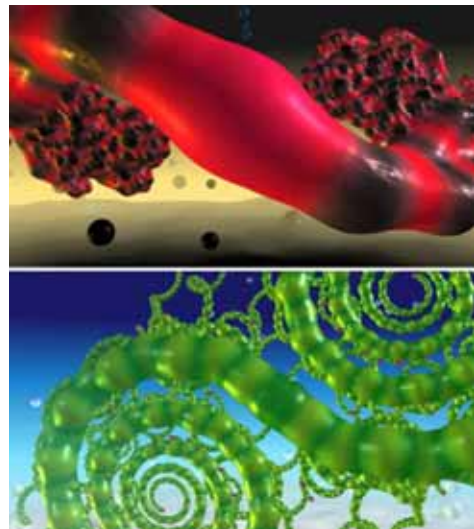


Figure 1: The GROWTH models.

Our research purpose is to build a new application of the GROWTH model on interactive arts that allows it to react emotionally to the participant's movements. The rendering methods presented in this paper is based on our previous researches. In the previous researches, we have proposed 1) the real-time rendering method for the GROWTH model with structural color [2] and 2) the method for animating the model interactively with participants' arm movement [3]. We will briefly introduce these methods in the following chapters. In the previous researches, the participant's gesture was reflected only to the shape and motion of the model. And interactions between the participant's motion and the surface property such as branch color were disregarded.

In this paper, we present an improvement of this model to capture and reflect tilt angle of participant's torso and the position of the whole body in the sensing area (in section 2). Also we present a new coloring method so-called "branch direction dependent color", and combined this new method with iridescent property caused by thin film interference into the model's surface characteristics (in section 3). We also propose a method for generating various color combinations each time the model develops its shape when the participant comes in front of the artwork (in section 4). These methods enable the model to express not only the participant's motion in terms of its movement but also the changes in its surface color. For example, if the participant moves intensely, the model shows colorful blinking appearance.

2. THE GROWTH MODEL GEOMETRY

The GROWTH model is described by the set of rules (scaling, twisting and ramification) between its root and the subsequent branch (Figure 2). The input parameters include the following:

- Height / radius / twist angle / tilt angle of the root.
(H_0, R_0, θ_0 and Φ_0)
- Height scaling factor / radius scaling factor / twist angle / tilt angle of the subsequent bough.

($H_1/H_0, R_1/R_0, \theta_1$ and Φ_1)

- Height scaling factor / radius scaling factor / twist angle / tilt angle of the subsequent branched twig.
($H_2/H_0, R_2/R_0, \theta_2$ and Φ_2)
- Ramification cycle.
- Threshold value of isosurface.

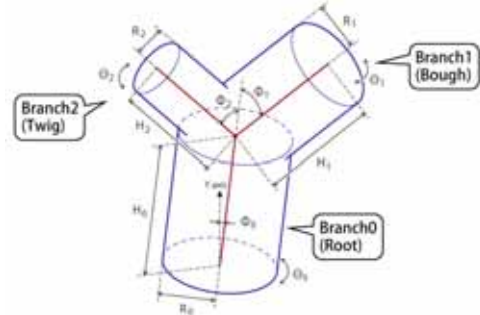


Figure 2: Branching structure of the GROWTH model.

The GROWTH model structure is computed starting from its root, recursively applying the rules of development and ramification with the above parameters. Also, as these parameters animate, the model organically changes its shape accordingly.

The model is rendered with "Metaballs" (geometric isosurface), which enable the model to take on the biotic flexibility. To achieve this characteristic, (1) We divided the space into grids, (2) integrated field values of all branches for each lattice point, and (3) the geometric isosurface was computed using the marching cubes algorithm. Furthermore, GPU computation was introduced in order to accelerate rendering speed.

2.1 Smoothing the joints of branches

In our previous researches, we put a straight, cylinder-shaped field for each branch. Therefore, the resulting model had a straight-lined impression in its appearance. In this paper, instead of the cylinder, we put three sphere-shaped fields for each branch. Furthermore, we offset the first/third sphere positions

according to the relative positions of previous/next branches in order to make the joint of the branches smoother and to obtain more life-like flexibility to the model (Figure 3).

The positions of three spheres $S_1(\vec{S}_1)$, $S_2(\vec{S}_2)$ and $S_3(\vec{S}_3)$ for the n-th branch between end points P_n and P_{n+1} were obtained by the following equations:

$$\begin{aligned}\vec{S}_0 &= \vec{Q}_n + \frac{1}{3}\vec{Q_nP_n} + \frac{1}{6}\vec{Q_nQ_{n-1}} \\ \vec{S}_1 &= \vec{Q}_n \\ \vec{S}_2 &= \vec{Q}_n + \frac{1}{3}\vec{Q_nP_{n+1}} + \frac{1}{6}\vec{Q_nQ_{n+1}}\end{aligned}\quad (1)$$

where $Q_n(\vec{Q}_n)$ is the midpoint of the n-th branch.

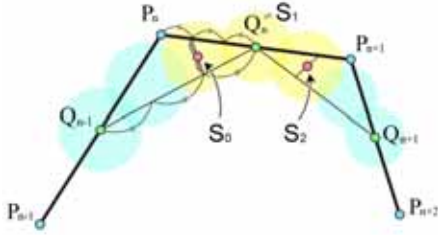


Figure 3: Defining the positions of three sphere-shaped fields for each branch.

3. TWO TYPES OF SHADING

We decided to set the surface color property as the evolving characteristic of the model, and changed it by the growth of the model. Two types of surface property were computed; 1) The structural color caused by thin film interference (Figure 4) and 2) the “branch direction dependent color” (Figure 5). These two properties were mixed according to the growing ratio of the model.

3.1 Making the structural color

Born and Wolf describes the spectrum calculation technique for thin film interference in their book [5]. Spectral reflectance of each reflection angle θ ($0^\circ \sim 90^\circ$) and each thickness of the thin film d ($0 \sim 1\mu\text{m}$) are obtained. Reflectance intensities are calculated in the 380nm to 710nm range, every 10nm apart. For the spectral intensity of incident light, we applied CIE standard

illuminant D65 [6].

The irradiance function is converted to the XYZ color intensity using CIE1964 color-matching function [7] and integrated through the range of wavelength, then converted to the RGB color values.

The structural color was computed in advance and passed to the vertex shader as two-dimensional texture. At each vertex of the model, texture coordinates was computed from the viewing angle of the surface and the pseudo thickness of the thin film d was computed by the following equation considering that the film becomes thick at the bottom by the effect of gravity;

$$d = (1 - k) \cdot \frac{1 - n_y}{2} + k \cdot \text{noise} \quad (2)$$

where n_y is the Y component of the normal vector, *noise* is Perlin noise factor and k is the contribution rate of the noise.

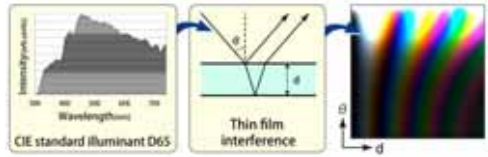


Figure 4: The structural color caused by thin film interference. Color changes depending on the thickness of thin film d and the reflection angle of light θ .

3.2 Making the branch direction dependent color

The branch direction dependent color was obtained by the following procedure: 1) Setting *direction feature vector* for each branch 2) weighted averaging the *branch direction field* values for each lattice point and then calculating *branch direction field* value for each vertex by linear interpolation, and 3) converting *branch direction field* value to the surface color. Each step is to be discussed in the following sub-chapters.

3.2.1 Setting direction feature vectors

When computing the branch structure of the model described in the section 2, the branches

were divided into eight categories depending on their directions. The division was done by simply checking the signs of X, Y and Z components of the branch direction. According to the category, the *direction feature vector* \vec{f} of i-th branch is as

$$\vec{f}_i = \begin{pmatrix} (b_{ix} \geq 0)? 1: -1 \\ (b_{iy} \geq 0)? 1: -1 \\ (b_{iz} \geq 0)? 1: -1 \end{pmatrix} \quad (3)$$

where b_{ix} , b_{iy} and b_{iz} are X, Y and Z components of the direction of i-th branch.

3.2.2 Weighted averaging the branch direction field values

Three dimensional vector value named “*branch direction field*” was introduced into the lattice points used for the modeling of metaballs geometry with marching cubes algorithm. For each lattice point, the *branch direction field* value \vec{g}_p was obtained by weighted averaging of the *direction feature vectors* of all branches as

$$\vec{g}_p = \frac{\sum_i k_{i,p} \vec{f}_i}{\sum_i k_{i,p}} \quad (4)$$

where $k_{i,p}$ is the value determined by the i-th branch’s radius r_i and the distance between the i-th branch and the lattice point $d_{i,p}$. The closer the i-th branch get, the larger influence rate of the *direction feature vector* of i-th branch. For $k_{i,p}$, we applied the same value with the density function used for modeling of the metaball geometry obtained by following function [8]:

$$k_{i,p} = \begin{cases} \left(1 - \left(\frac{d_{i,p}}{r_i}\right)^2\right)^2, & \text{if } d_{i,p} < r_i \\ 0, & \text{if } d_{i,p} \geq r_i \end{cases} \quad (5)$$

The *branch direction field* value for each vertex of the model was computed by linear interpolation of the *branch direction field* values of neighbour lattice points. The interpolation was done at the same time and in the same manner with the interpolation of the density in the marching cubes method for modeling of the metaball geometry.

3.2.3 Converting to the surface color

From the calculations above, the *branch direction field* value was obtained for each vertex of the model (\vec{g}_v). Range of the X, Y and Z components of \vec{g}_v is -1 to 1. This value was passed to vertex shader and used for determining the branch direction dependent color at each vertex.

We introduced eight *palette colors* ($C_1 \sim C_8$). The branch direction dependent color at the vertex v was obtained as

$$\begin{aligned} C_v = & (1 - ratio_x)(1 - ratio_y)(1 - ratio_z)C_1 \\ & + ratio_x (1 - ratio_y)(1 - ratio_z)C_2 \\ & + (1 - ratio_x) ratio_y (1 - ratio_z)C_3 \\ & + ratio_x ratio_y (1 - ratio_z)C_4 \\ & + (1 - ratio_x)(1 - ratio_y) ratio_z C_5 \\ & + ratio_x (1 - ratio_y) ratio_z C_6 \\ & + (1 - ratio_x) ratio_y ratio_z C_7 \\ & + ratio_x ratio_y ratio_z C_8 \end{aligned} \quad (6)$$

where $ratio_x, ratio_y$ and $ratio_z$ are the interpolation ratio of *palette colors* and are calculated as

$$ratio_{x,y,z} = \begin{cases} 0, & \text{if } \vec{g}_{v,x,y,z} \leq -\frac{a}{2} \\ \frac{1}{a}\vec{g}_{v,x,y,z} + 0.5, & \text{if } \frac{a}{2} < \vec{g}_{v,x,y,z} < \frac{a}{2} \\ 1, & \text{if } \frac{a}{2} \leq \vec{g}_{v,x,y,z} \end{cases} \quad (7)$$

where a is the parameter to determine the smoothness of the color shift between branches with different colors.

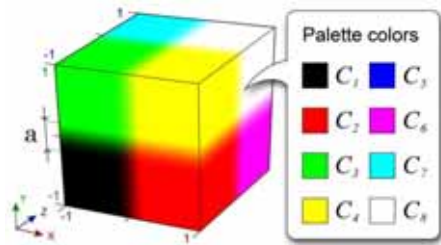


Figure 5: The branch direction dependent color. This figure shows how the influence rate of the palette color changes depending on the branch direction field value.

4. INTERACTION WITH PARTICIPANT

We decided to add the following three kinds of interactions between the model and the participant with different time scales:

- (1) Short term. The model changes its shape dynamically according to the participants' movement.
- (2) Middle term. The model grows/shrinks by using the information of existence of the participant as a trigger. i.e. The model grows when the participant is in front of the artwork, and starts to decay when the participant exits.
- (3) Long term. The model displays various different colors each time the model develops its shape when the participant come in front of the artwork.

4.1 The short term interactivity

In order to add the short term interactivity, the participants' shoulder and elbow angles were reflected to the angle of branches as shown in Figure 6. The participant's movement of the arms and body was captured by the Kinect sensor in the form of position data of each joint. The GROWTH model shape was generated using the angles of the arms as the parameter.

For each branch ramified from the initial root, different values are applied as follows:

The initial root

$$\theta_0 = \theta_{torso}, \Phi_1 = \phi_{torso}$$

Branch1 (Bough, right branch)

(if $g \pmod{2} = 0$)

$$\theta_1 = \theta_{R,sho}, \Phi_1 = \phi_{R,sho}$$

$$\theta_2 = \theta_{L,sho}, \Phi_2 = \phi_{L,sho}$$

(if $g \pmod{2} = 1$)

$$\theta_1 = \theta_{R,elb}, \Phi_1 = \phi_{R,elb}$$

$$\theta_2 = \theta_{L,elb}, \Phi_2 = \phi_{L,elb}$$

Branch2 (Twig, left branch)

(if $g \pmod{2} = 0$)

$$\theta_1 = \theta_{L,sho}, \Phi_1 = \phi_{L,sho}$$

$$\theta_2 = \theta_{R,sho}, \Phi_2 = \phi_{R,sho}$$

(if $g \pmod{2} = 1$)

$$\theta_1 = \theta_{L,elb}, \Phi_1 = \phi_{L,elb}$$

$$\theta_2 = \theta_{R,elb}, \Phi_2 = \phi_{R,elb} \quad (8)$$

where θ_{torso} and ϕ_{torso} are twist and bend angles of participant's torso,

$\theta_{[R,L],[sho,elb]}$ and $\phi_{[R,L],[sho,elb]}$ are twist (bend) angles of the shoulder/elbow joint of participant's right/left arm respectively, and g is the generation of ramification starting from 0.

In addition, we related the distance between the participant and the Kinect sensor to the distance between the root position of the model and the camera, so that the GROWTH model behaves like a "mirror image" of the participant.

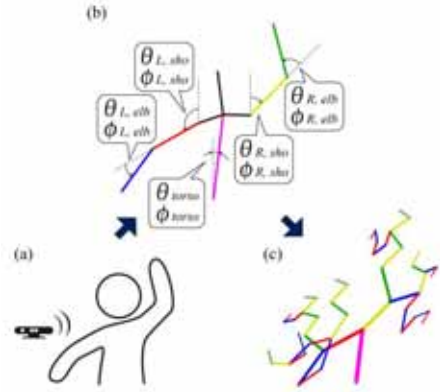


Figure 6: Adding the short term interactivity.

(a) Participant's gesture is captured with Kinect sensor.

(b) Angles of the torso and arm joints are calculated.

(c) The angles are applied to the branch structure.

4.2 The middle term interactivity

In order to add middle term interactivity to the model, we decided to assume the period while a participant is in front of the artwork as the "lifetime" of individual model, and made the model grow and shrink through the lifetime (Figure 7).

The max branch length of the model is calculated in the first place. Within the limit of such length, the variable branch length is increased / decreased by using the information of participant's existence as a trigger. And for each frame, the branch structure is calculated only in the length of branch length from the root.

While a participant is in front of the artwork (purple line), the branch length (orange graph) increases to the maximum branch length. And

when the participant exits the artwork, the model starts to shrink to the minimum branch length.

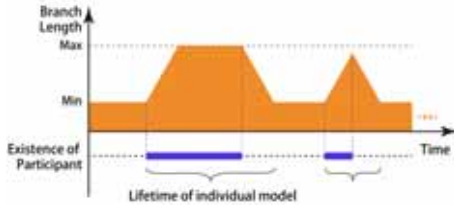


Figure 7: The middle term interactivity.

4.3 The long term interactivity

In order to add the long term interactivity, we decided to set the surface color property as the evolving characteristic of the model, and change it by each “generation”.

We applied the structural color caused by thin film interference in the “seed” state of the model and made the individual colored “branch direction dependent color” pattern emerges as the model grows.

When the model shrinks (i.e. the *branch length* value reached the minimum), the *palette colors* are re-computed and passed to the fragment shader in preparation for the next generation of the model. In order to increase the visual impact, the palette colors are randomly picked from HSV color space avoiding the medium saturation and brightness which causes the dull color.

According to the growth of the model, the surface characteristic is smoothly changed from the structural color (*Structural*) to the “branch direction dependent” color (*BDD*). Mixing two colors, the surface color *Out* was calculated as follows:

$$\begin{aligned} \text{Ratio} &= g_ratio \times (1.0 - \text{Base}), \\ \text{Out} &= \text{Structural} \times (1.0 - \text{Ratio}) + \text{BDD} \times \text{Ratio} \end{aligned} \quad (9)$$

where *g_ratio* is the growth ratio of the model which changes depending on the *branch length*, and *Base* is a contribution ratio of the structural color texture when the model grows to the max branch length.

5. RESULT AND FUTURE WORK

Figure 8 shows the rendered results with captured images and tracked skeleton by Kinect. Parameters used for each result are listed in Table 1. Note the short and long term interactivities - the model changes its shape according to the participant’s posture and the various color combinations appear for each example of the model. Also note that the color regularity appears at the spiral part of the model. Result #4 shows that the model becomes larger as a participant comes close to the sensor. Figure 9 shows the middle term interaction of the model. Note that the “branch direction dependent” color pattern appears according to the growth of the model.

Figure 10 shows the result of smoothing for the joints of branches discussed in section 2.1. Note that compared to the cylinder-shaped field (previous approach), three sphere-shaped fields (new approach) configures smoother connection at the joints and therefore the model becomes more rounded and flexible in appearance.

The prototype implementation was written in C++, using OpenGL and CUDA. All tests were conducted on a PC with an Intel Core i7 3.4GHz Processor and an NVIDIA GeForce GTX780TI graphics card. Each scene was captured at around 27 fps.

For future improvement on this model, we would like to make one level of interaction affecting the other two levels. For example, the more changes in movement of the model, the more the color of the next generation’s model becomes vivid.

6. CONCLUSIONS

We presented the new method for generating geometry of the GROWTH model for an interactive artwork. We captured the human posture data including the angle of arm joints, tilt of torso and the position of the whole body in the sensing area and related such data to the parameters determining the model geometry. We also introduced the new coloring method so-called “branch direction dependent color” for model’s surface property.

As a result, the model changes its shape according to the participant's movement in practically real-time, and as the model changes its shape, causing each branch to change its direction, the model shows a colorful blinking appearance. In conclusion we could gain a new perspective for improving the model to the interactive artwork.

REFERENCES

- [1] Y. Kawaguchi. A morphological study of the form of nature. In *Proceedings of ACM SIGGRAPH '82*, pages 223-232. ACM Press, 1982.
- [2] S. Iwasawa and Y. Kawaguchi. Rendering Methods for the GROWTH Model with Structural Color. In *Proceedings of ASIAGRAPH '13*, pages 27-30. 2013.
- [3] S. Iwasawa and Y. Kawaguchi. Animating the GROWTH Model Using Human Motion Data. In *Proceedings of ASIAGRAPH Forum in Kagoshima*, pages 11-14. 2013.
- [4] S. Iwasawa and Y. Kawaguchi. Color Designing for GROWTH Model with Lifelike Interaction. In *Proceedings of the 11th annual conference of Asia Digital Art and Design Association*, pages 68-71. 2013.
- [5] M. Born and E. Wolf. Principles of Optics I, pages 91-93. Tokai University Press (Japanese edition). 1974.
- [6] CIE D65 Standard Illuminant Data. In the website of the *Munsell Color Science Laboratory*, Rochester Institute of Technology. <http://www.cis.rit.edu/research/mcsl/online/cie.php>
- [7] CIE 1964 10-deg, XYZ Colour Matching Functions. In the website of the *Colour & Vision Research laboratory and database*, <http://www.cvrl.org/cie.htm>
- [8] S. Murakami and H. Ichihara. On a 3D display method by metaball technique. In *Transactions of the Institute of Electronics, Information and Communication Engineers* (In Japanese). J70-D(8): pages 1607-1615. 1987.

Table 1: Parameters used for the rendering results in Figure 8.

Common parameters				
voxel size	128 x 128 x 64			
H_1/H_0 and R_1/R_0	0.9			
H_2/H_0 and R_2/R_0	0.7			
Other parameters				
	#1	#2	#3	#4
Threshold of isosurface	0.448	0.415	0.487	0.61
θ_{torso} (degrees)	-16.18	-6.57	-0.04	12.82
ϕ_{torso}	7.60	17.67	21.19	27.70
$\theta_{R,sho}$	0.57	16.64	-0.55	-13.02
$\phi_{R,sho}$	102.4	64.58	90.15	144.4
$\theta_{R,elb}$	-172.5	153.9	171.2	125.7
$\phi_{R,elb}$	63.39	70.60	25.67	24.06
$\theta_{L,sho}$	164.1	116.5	33.76	35.45
$\phi_{L,sho}$	120.9	46.87	32.81	58.88
$\theta_{L,elb}$	-118.9	-130.7	-46.64	-37.68
$\phi_{L,elb}$	40.85	98.58	35.62	28.03
Rendering status				
Occupied voxels (%)	35758	35228	28249	41903
	3.41%	3.36%	2.69%	4.00%
# of vertices	214788	212604	169836	252036
fps	27.0	27.4	27.5	27.1

ACKNOWLEDGMENTS

This work was supported by JSPS KAKENHI Grant Number 26540098 (Grant-in-Aid for Challenging Exploratory Research).

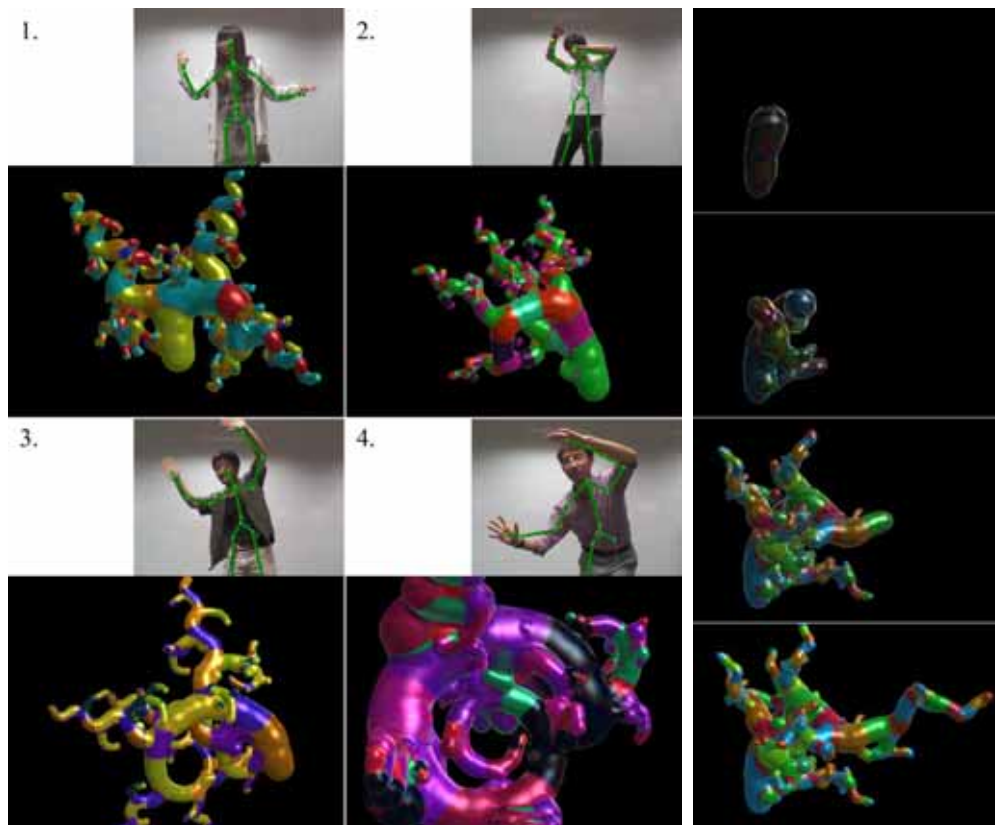
ABOUT THE AUTHORS

1. Shun Iwasawa : Assistant Professor of Interfaculty Initiative in Information Studies, The University of Tokyo.

Mail: shun.iwasawa@iii.u-tokyo.ac.jp

2. Yoichiro Kawaguchi : CG Artist, Professor of The University of Tokyo.

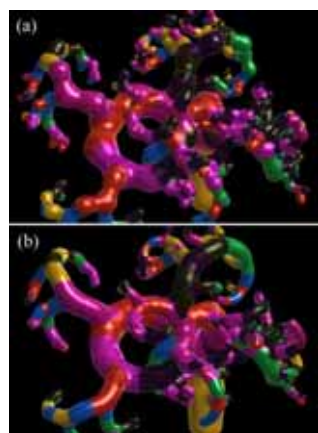
Mail: yoichiro@iii.u-tokyo.ac.jp



[Upper-left] Figure 8: Rendered results with various gestures. Parameters are listed in Table 1. For each result, (top) captured images by Kinect sensor with tracked skeleton by Kinect NUI Library, and (bottom) rendered results.

[Upper-right] Figure 9: The middle term interaction. With no participant in front of the artwork, the model is in the “seed” state with the structural color (top). When the participant comes, the model grows and changes its surface to the individual colored “branch direction dependent” pattern (top to bottom).

[Lower-right] Figure 10: Smoothing the joints of branches (section 2.1). Each result is rendered by using (a) the straight, cylinder-shaped fields and (b) the three sphere-shaped fields with offset.



THE GEOMETRIC AND COMPUTER MODELING SHAPING TECHNICAL SURFACES

Vladimir VOLKOV¹ and Aleksey LYASHKOV²

¹Siberian State Automobile and Highway Academy, Russia

²Omsk State Technical University, Russia

ABSTRACT: The process of shaping of technical surfaces is represent interaction tool (forming tool) and detail (form shapes elemental or metal block) in their relative movements. Found that the main targets of shaping are: 1) the discriminant family of surfaces formed by the movement of shape-generating elements relative to the work piece; 2) enveloping - model of the real surface of the part obtained after processing, including undercuts and transition curves on separate portions of the profile; 3) models in the process of shaping the cut layers. When modeling shaping of these objects a lot of poorly resolved or unresolved issues that form a single scientific problem-the problem of qualitative forming the surface of the tool and the work piece, surface generated by this tool. Improvement of the famous metal-cutting tools, intensive development of computer-aided design systems, requires further improving the methods of shaping conjugate surfaces. In connection with this important role belongs to the study of the processes of formation of technical surfaces with positive aspects of analytical and numerical methods and techniques associated with the use of geometric and computer modeling. In this paper we formulated and solved the problem of developing a mathematical, geometric and algorithmic support systems computer-aided design of the cutting tool based on computer simulation modeling of the process of shaping surfaces.

Keywords: Geometric modeling, simulation, formation, discriminant, bevel helical surface.

Manufacture of products in several branches of industry bound respective with technological processes forming geometrically complex parts surfaces. These details include the rotors of screw pumps, impellers turbines, compressors and pumps, impeller fans, compressors and others.

An important role is given to the issues of shaping during the design of the cutting tool. One of the elements of the design process is to design surface the tool's. The solution to this problem is considered in many papers. In many of them to perform the calculation is necessary output the appropriate dependencies for the different source data. This uses the classical method of differential geometry or kinematic method [1]. One of the new approaches proposed in [2] and to some extent used in

[3-7]. Often the computational dependencies have the shape of of transcendental equations. To solve them, numerical methods are used. All this complicates the process of profiling tool.

Definition enveloping family of surfaces be based on development of mathematical models of the shaping process. This approach does not use differential methods.

However, the effective solution of the considered problems formation of complex surfaces can be made with the use of methods of geometrical simulation by means of computer graphics and CAD systems [8, 9]. Modern computer technology, providing the possibility of modeling the shaping process, allow us to determine the influence of various parameters of the tool, as well as its location on the shape of the work piece surface. In addition, they

provide the opportunity to develop software, allowing to model the motion of morphogenesis in the automated mode. In this case, the problem solved shaping performed with necessary accuracy [10].

It provides a solution to problems of technological software sizing different surfaces:

- roll forming part of the cutting tool;
- appointment conditions for reciprocal movement of the tool and the work piece;
- forecasting possible errors shaping surface of the work piece and other tasks.

It is proposed to study the processes of formation of technical surfaces performed using analytical and numerical methods and techniques associated with the use of geometric and computer modeling. Uniting them is to study maps the orthogonal projection of two-dimensional surfaces and three-dimensional hypersurface given by an implicit and parametric equations on the coordinate plane and hyperplane.

Let the original surface, defined by the equation in implicit form:

$$F(x, y, z) = 0 \quad (1)$$

At points criminant this surface relative to the coordinate plane XY is satisfied:

$$F_z(x, y, z) = 0. \quad (2)$$

Considering equation (2) as an equation of another surface and determining its intersection with the original surface (1), we obtain criminant surface (1) with respect to the coordinate plane XY. Criminant this projection on the coordinate plane XY is the discriminant of the surface (1) and the envelope of a family of plane curves congruent if the surface (1) we obtain a map of these curves in the space R^3 .

Found that equation (2), reflecting the tangency condition "vertical" plane with a given surface can also be seen as a necessary condition for the existence of conditional extremum function whose graph curves are obtained at the intersection of the surface (1) planes parallel to the coordinate planes XZ or YZ.

To determine the sufficient conditions for

the existence of extrema on the curves defined considered the second differential of the Lagrange:

$$d^2L = \square \frac{F_{zz}}{F_y} \square dz^2 \quad (3)$$

Then from (3) that if $\frac{F_{zz}}{F_y} > 0$, the point of the

test section of the surface plane parallel to the coordinate plane XZ or YZ is conditional maximum point; if the corresponding point - the point of a conditional minimum. $\frac{F_{zz}}{F_y} < 0$,

for further studies.

As a result criminant D is the union of the set of extreme points, namely

$$D = \sum_{i=1}^n \min f(x, z) \square \max f(x, z) \Big|_{x=a_i}.$$

Variable is the z coordinate in its domain.

The analysis of the total curvature surface possible to establish the following:

a) if $F_{zz} = 0$, and $F_y [F_{zx} \square F_x [F_{zy} \square 0$ then the Gaussian curvature at the surface is negative kriminatnoy line and the point on the surface is hyperbolic, the mean curvature of the surface it is directly proportional to the value of the curvature of the curve obtained at the intersection of the plane parallel to the XY;

b) if $F_{zz} \square 0, |F_x| + |F_y| \square 0$, a curved surface on a plane (sectional plane perpendicular to the Z axis has an inflection point and the corresponding point kriminantnoy line hyperbolic surface;

c) if $F_{zz} = 0$ and $F_y [F_{zx} \square F_x [F_{zy} = 0$ then the corresponding point of the parabolic surface;

d) if $F_y^2 [F_{xx} \square 2F_x [F_y [F_{xy} + F_x^2 [F_{yy} = 0$ and $F_y [F_{zx} \square F_x [F_{zy} = 0$ then the cor-

responding point of the parabolic surface;

e) if $F_x = F_y = 0$, as is the special point on the original curve, and on the surface, as well as its discriminant;

f) if $2F_x F_y F_{xy} F_{xx} F_y^2 F_{yy} F_x^2 \neq 0$, $F_{zz} \neq 0$ and $F_y [F_{zx} F_x [F_{zy} \neq 0]$, the corresponding point of the surface may be either elliptical or hyperbolic.

By analogy with the previous problem investigated display orthogonal projection of the surface given by the parametric equations:

$$x = f_1(u, v), y = f_2(u, v), z = f_3(u, v).$$

In this case the discriminant D is the union of the set of extreme points, namely:

$$D = \sum_{i=1}^n \min f_1(u, v) \neq \max f_1(u, v) \Big|_{f_2(u, v)=a_i, z=f_3(u, v)}.$$

Variable is one of the parameters of the surface in its definition domain.

The total curvature of the surface at the points of its line will discriminant

$$K = \frac{(a[f_{1uv} \neq b[f_{2uv}]) (a[f_{1uv} \neq b[f_{2uv}]) \neq (a[f_{1uv} \neq b[f_{2uv}])^2}{(E[G \neq F^2])^2},$$

where $a = f_{2u} [f_{3v} \neq f_{3u} [f_{2v}$,

$$b = f_{1u} [f_{3v} \neq f_{3u} [f_{1v}.$$

$$E = f_{1u}^2 + f_{2u}^2 + f_{3u}^2,$$

$$G = f_{1v}^2 + f_{2v}^2 + f_{3v}^2,$$

$$F = f_{1u} f_{1v} + f_{2u} f_{2v} + f_{3u} f_{3v},$$

Analyzing the expression, you can set the following:

a) If $(a[f_{1uv} \neq b[f_{2uv}]) (a[f_{1v} \neq b[f_{2v}]) = 0$, then the Gaussian curvature of the surface at the points discriminant negative and the point on the surface is hyperbolic;

b) where $r(u, v) \neq C2$, then the corresponding point of the parabolic surface;

c) if $r_u \times r_v = 0$, it is a particular point as the original curve, and on the surface, as well as its sketch;

$$d) \text{ if } (a[f_{1uv} \neq b[f_{2uv}]) (a[f_{1v} \neq b[f_{2v}]) \neq (a[f_{1uv} \neq b[f_{2uv}])^2 > 0$$

the corresponding point on the surface - the elliptical.

The obtained results are extended to the definition of the envelope parameter family of surfaces. Initially we investigate the three-dimensional mapping of the hypersurface given in four-dimensional equation in implicit form:

$$F(x, y, z, t) = 0$$

For her discriminant D is the union of the following set of extreme points:

$$D = \sum_{j=1}^n \left(\sum_{i=1}^n \min f(x, y, t) \Big|_{x=x_i} \neq \max f(x, y, t) \Big|_{x=x_i} \right)_{y=y_j},$$

where variable is the coordinate t.

Let now consider the three-dimensional hypersurface in four-dimensional space defined by the parametric equations:

$$x = f_1(u, v, \Delta),$$

$$y = f_2(u, v, \Delta),$$

$$z = f_3(u, v, \Delta),$$

$$t = f_4(u, v, \Delta).$$

Then discriminant D surface is the union of the set of extreme points, namely

$$D = \sum_{j=1}^n \left(\sum_{i=1}^n \min f_1(u, v, \Delta) \Big|_{y=y_i} \neq \max f_1(u, v, \Delta) \Big|_{y=y_i} \right)_{z=z_j}.$$

Variable is one of the parameters of a hypersurface in its domain.

Thus, the obtained results allow us to determine the discriminant of two-dimensional surface and three-dimensional hypersurface, and hence the envelope of a family of plane curves and surfaces with one voice. Analytically - through conditional extrema corresponding functions, based on the method of Lagrange multipliers, and computational methods - without the use of

differential surface parameters.

In applied problems one-parameter family of congruent curves often associated with the centroid of which rolls without slipping on a fixed centroid. Let be a family formed curve associated with the circle, rolling without slipping along a straight line. If the original curve is given by the parametric equations $x = x(t), y = y(t)$, is formed by a family in a fixed coordinate system OXY be determined as

$$x = x(t) [\cos \varphi] + y(t) [\sin \varphi] + R [\varphi],$$

$$y = -x(t) [\sin \varphi] + y(t) [\cos \varphi],$$

where φ - angle of rotation around the center of the curve forming the circle, and R - the radius of this circle.

Performing the mapping of the family of congruent curves in space R^3 , we obtain in the coordinate system XYZ (Z-axis together with the X and Y axes form a right-handed system) surface Ψ , defined by the equations:

$$x^\varphi = x(t) [\cos \varphi] + y(t) [\sin \varphi] + R [\varphi],$$

$$y^\varphi = -x(t) [\sin \varphi] + y(t) [\cos \varphi],$$

$$z^\varphi = p' [\varphi],$$

(4)

where p' - some constant.

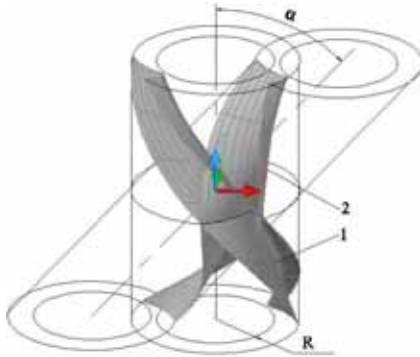


Fig. 1: A polygonal model HS-1 and ISS-2; R is the radius of the cylinder and is equal to the radius of the bending circle

Associate with the fixed coordinate system XYZ left cylindrical helical surface HS (Fig. 1)

formed a helical motion of the original curve.

Its equations are:

$$x^\varphi = x(t) [\cos \varphi] + y(t) [\sin \varphi],$$

$$y^\varphi = -x(t) [\sin \varphi] + y(t) [\cos \varphi],$$

$$z^\varphi = p [\varphi],$$

(5)

where p - parameter screw motion, and $p = H / 2 [\varphi]$, and H - pitch of the helical surface.

Comparing equations (4) and (5), taking $p' = p$

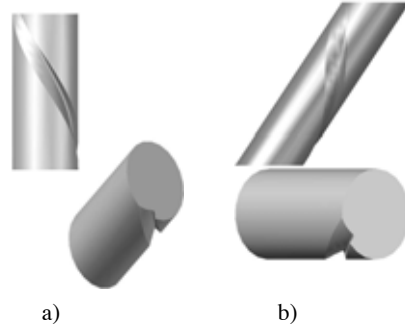


Figure 2: Solid model: a) the body with the HS and its oblique projection; b) the body with the ISS and its orthogonal projection

and $R [\varphi] = \frac{R}{p} [\varphi]$, we obtain:

$$x^\varphi = x^\varphi + \frac{R}{p} [\varphi],$$

$$y^\varphi = y^\varphi, \quad (6)$$

$$z^\varphi = z^\varphi.$$

Considering the transition from HS and Ψ to the surface as a result of the geometric transformation is easy to see that (6) describes an affine transformation, since is linear. The thus converting the surface (4) (Fig. 1) is called

the inclined helical surface (ISS) [11].

Research maps HS oblique projection, and ISS - orthogonal projection onto the coordinate plane shown their coincidence (Fig. 2). This result establishes a relationship of some planar and spatial profiling tasks mating surfaces.

Surface models are used to solve various problems of morphogenesis. Since the auxiliary surface [11], derived from the family plane curves used in CAD-systems for the qualitative characteristics of the envelope, as well as for studies of its forms depending on the radius of the centroids that illustrates figures 3 and 4.

The next step in the application of computer technology problems forming surfaces is to develop solid models of parts with geometric shape than the 3D primitives used in known CAD. For these purposes may be used, in particular, surface models developed previously. Thus, Fig. 5 shows a model of a helical groove parts. It is derived from the blank cutting off body parts inclined helical surface. At the same technology to build a solid model based impeller models and allowances removed (Fig. 6). The very same solid model billet impeller obtained using standard means of most CAD systems. To develop the most effective schemes interblade volume stock removal is necessary to analyze the geometric model of the body of this allowance - P. The proposed approach to the selection of alternatives for the disposal of most of the allowance laid geometric methods of synthesis and the capabilities of modern hardware and software for three-dimensional bodies morphogenesis [12].

Selection step rational schemes remove most of the allowance interblade volume involves the following tasks:

- a) analysis of the shape of the surface bounding the volume under investigation;
- b) removing a portion of the allowance interblade volume using geometric modeling operations;
- c) the choice of a rational process step, or combinations thereof, removing parts of the allow-

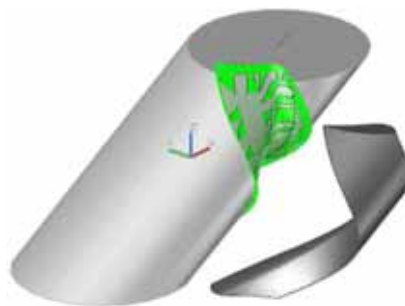


Fig. 5: Forming part model with inclined screw surface.

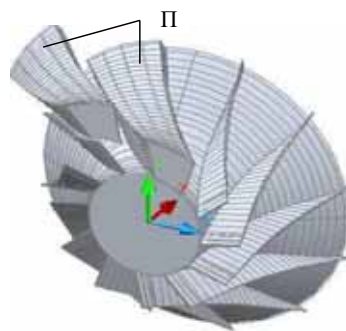


Fig. 6. The solid model of the impeller;
P - removed interscapular allowance

ance based on analysis of the performance of each of them;

d) selection process, equipment and tools for dimensional pretreatment impeller.

COMPUTER SIMULATION OF SHAPING

The simulation results from the previous steps are used to develop algorithms and programs for the appointment of the necessary technological conditions shaping the details rational methods. for this purpose use known solid-state modeling operations. So kinematic operation is used to specify multitude provisions of the instrument in the process of shap-

ing details, position setting and the orientation of the tool can be automated using programming languages. This operation allows to model the most complex schemes of education is difficult to profile surfaces. Further, using Boolean operations, is formed required surface .

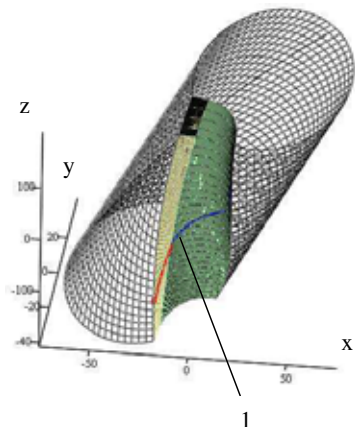


Fig. 3: A polygonal model ISS general position; 1 - criminant surface with respect to the XY plane

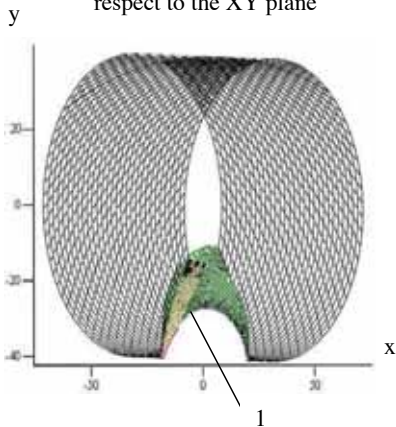


Fig. 4: Orthogonal projection of ISS on the plane XY; 1 - discriminant surface (envelope of a family of plane curves)

In the tasks of forming surfaces of the parts to be processed method centroid and without center rounding effectively used the capabilities of modern CAD-systems and related programming languages. One example may be the solution of direct and inverse problems of modeling of surfaces: helicoid - rotational surface - helicoid.

In solving the direct problem of initial data are body model with a spiral groove and workpiece model as section of the cylinder. Cylinder size and location determined by the parameters setting shaping element relative to the workpiece. Solid models of initial data shown in figure 7. Modeling of surface formation of rotation on the created model and the installation parameters is performed in accordance with the developed algorithms and programs

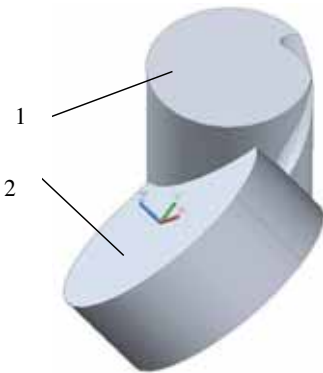


Fig. 7. Mutual position shape-generating element - 1 (HS) and model workpiece-2 for solids of revolution

(fig. 8).

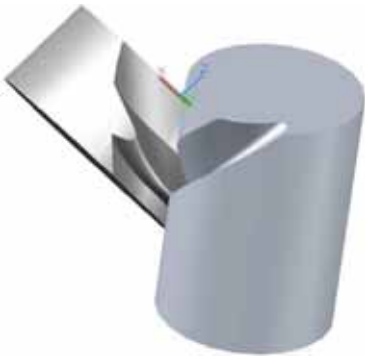


Fig. 8. Forming a solid of revolution on the model of the body with the screw surface

Due to the fact that the surface of the body of revolution are uniquely determined by its axial cross-section, the simulation is subjected to forming only a part of the material. Fragments of the rotation body and its axial section shown in Fig. 9. The resulting axial section may be subjected to editing as to replace profile technological curves and for some other reasons. This section is used to obtain a solid model of a body of revolution on which the model is created disk cutter and its drawing. Figure 10 illustrates the effect of the angle between the axes of the workpiece and the screw on the surface shape of the profile cutters.

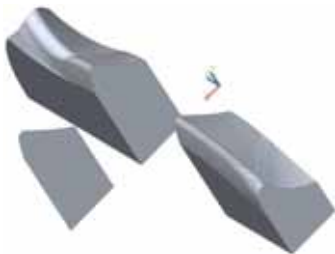
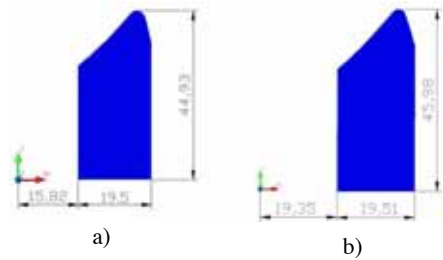


Fig. 9. Fragments of a body of revolution, intersect a coordinate plane, and its axial section



1. The angle between the axes 49° (a) and 47° (b).
2. The rotation angle of the surface 93° .

Fig. 10: Axial profiles of rotation of a body for various angles between the axes models

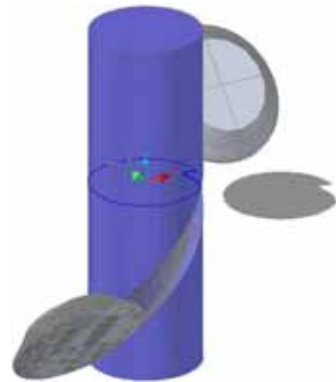


Fig. 11. The result of simulation helical groove

Built on the solid model of the axial section of the tool used to solve the inverse problem of morphogenesis. In this case, the simulation result of forming a body with a spiral groove (Fig. 11). The same figure shows a model of its end section which is used for comparative analysis with the original face profile. If the resulting deflection is within the permissible range, the modeling process is completed. Otherwise, this may change as the parameters of mutual arrangement of the tool models and products, and editing the profile cutter to produce the desired result.



Fig. 12. Computer modeling of the cut layers

Along with solving direct and inverse problems of shaping the developed algorithms and programs allow you to receive and solid models of the cut layers to select the optimal technological parameters of cutting

CONCLUSIONS

□ On the basis of the above stated use of geometric modeling and computer technologies, you can perform the following tasks:

□ development of surface models based on a family of plane curves, to help you visualize how discrete and in animation mode change the shape of the envelope according to the shape of the original curve and the of the radius centroid;

□ receive solid model of a part to select the method processing technique corresponding to the greatest volume removal allowance during pre-treatment;

□ provides the possibility of modeling the shaping process in an automated way;

□ appointment installation conditions, securing and relative movement of the forming tool and the machined surfaces;

prediction of conditions preventing technological losses of products during the production process, associated with the failure of the relative positions and the relative movement of work-piece and tools.

REFERENCES

- [1] F. L. Litvin. Alfonso Fuentes Geometry and Applied Theory. Cambridge University Press, 2004. – 816 p.
- [2] R. Thom. Sur la theorie des envelopes. J. de math. pur et appl. – 1962. – Vol. 41. – № 2. – PP. 177-192.
- [3] J.W. Bruce, and P.G. Giblin. Curves and features. Publisher "World" – 1988. – 263 p.
- [4] A. Liashkov, and V. Volkov. Aspects of display of surface projections. International Conference of Engineering Education Silesian University of Technology. – Gliwice, Poland, 2005. – PP.830-833.
- [5] A. Lyashkov, and V. Volkov. Mapping hypersurface orthogonal projection onto a hyperplane. Vestnik of the Irkutsk State Technical University. – 2012. – № 2. – PP. 18-22.
- [6] H. Pottmann and M. Peternell. Envelopes-computational theory and applications. In: Spring Conference on Computer Graphics, PP. 3–23. Comenius University, Bratislava (2000).
- [7] T. Schulz, B. Juttler. Envelope Computation by Approximate Implicitization – Industrial Geometry, August 2010 – 21p. – http://www.industrial-geometry.at/uploads/nm_report_106.pdf
- [8] T. Nikolaos. CAD-Based Calculation of Cutting Force Components in Gear Hobbing / T. Nikolaos, A. Aristomenis // Journal of Manufacturing Science and Engineering JUNE. – 2012. – Vol. 134.
- [9] A. Lyashkov. Shaping of harts with a helical surface by means of a disk mill // Russian Engineering Research, Allerton Press, Inc. – 2012. – Vol. 32, № 4. – PP.404-406.
- [10] A. Lyashkov. Simulation of shaping helical surfaces rack-type tool and worm hob. Metalworking. – 2011. – № 1(61). –PP. 2-7.

- [11] A. Lyashkov, Ju. Kaneva. Auxiliary surface in the simulation of shaping surface details with the help of computer graphics. Vestnik of the Kuzbass State Technical University – 2011. – № 5 (87). – PP. 75-80.
- [12] A. Liashkov, Ju. Vivdenko, A. Shutov, and S. Balanovski. Simulation of shaping complex surfaces detail. Metalworking– 2010. – № 4. – PP. 36-42.

ABOUT THE AUTHORS

1. Vladimir Volkov - Dr.eng.sc., professor. Head of Dep. of Descriptive geometry and Computer Graphics Federal State Budget Educational Institution of Higher Vocational Training "The Siberian Automobile and Highway Academy (SibADI)". e-mail: volkov_vy39@mail.ru postal address: Dep. of Descriptive geometry and Computer Graphics Federal State Budget Educational Institution of Higher Vocational Training "The Siberian Automobile and Highway Academy (SibADI)", RF, Omsk, 644008, pr. Mira 5.

2. Aleksey Lyashkov - Dr.eng.sc., associate professor of Dep. of Engineering Graphics Federal State Budget Educational Institution of Higher Vocational Training "Omsk State Technical University". e-mail: 3dogibmod@mail.ru

postal address: Dep. of Engineering Graphics Federal State Budget Educational Institution of Higher Vocational Training "Omsk State Technical University", RF, Omsk, 644008, pr. Mira

GEOMETRIC OPTIMIZATION IN MINKOWSKI SPACE

Bernhard Blaschitz

Wagner Biro Stahlbau, Austria

ABSTRACT:

We will study the set of circles in the plane in a *point set model*: Every circle is assigned to a point in $\mathbb{R}^{2,1}$ such that the first two coordinates are its center and the third is its radius. Every curve $l: \mathbf{p}(t) = (p_1, p_2, p_3)(t)$, $\dot{\mathbf{p}}(t) \neq 0$ in $\mathbb{R}^{2,1}$ defines a 1-parameter set of circles, their envelope is the intersection of a torsal surface of constant slope Γ_l with the xy -plane. This intersection is a curve and usually consists of two branches, which are real for $\langle \dot{\mathbf{p}}, \dot{\mathbf{p}} \rangle_L \geq 0$ and do not coincide for $\langle \dot{\mathbf{p}}, \dot{\mathbf{p}} \rangle_L > 0$, where $\langle \mathbf{x}, \mathbf{y} \rangle_L = x_1y_1 + x_2y_2 - x_3y_3$ is the pseudo Euclidean inner product in $\mathbb{R}^{2,1}$. Therefore, the images of curves, whose derivative is steeper than $\frac{\pi}{4}$, are not real.

Assuming a discrete set of circles \mathbf{p}_i in the plane, a real envelope is looked for. There exist different heuristics in the literature, e.g. the two common tangents of two neighboring circles, which results in envelopes consisting of circular arcs and line segments; their connection is C^1 at best.

Another approach is to fit an arbitrarily smooth B-Spline \mathbf{b} to the images of the \mathbf{p}_i in $\mathbb{R}^{2,1}$ via a quadratic optimization. That does not ensure that $\langle \dot{\mathbf{b}}, \dot{\mathbf{b}} \rangle_L > 0$, e.g. that the envelope is real, so some authors use a Hermite interpolation with flat curve segments or biarcs.

The new approach of this work is reformulating the original problem as a constrained optimization: the quadratic objective function minimizes the distance between the cyclographic images of circles \mathbf{p}_i and a cubic B-Spline \mathbf{b} by observing the footpoint problem, i.e. the curve is reparametrized in an iterative procedure, which brings a better fit.

The reality of the envelope results in a quadratic, but non convex constraint, which can be linearized. This linearization is discussed in detail, as its formulation is central to this work.

The problems discussed for circles are also generalized for spheres; in the 1-parameter case that leads to a new method for interpolation points in $\mathbb{R}^{3,1}$ by curves, which translates to interpolation of spheres by canal surfaces. Approximating 2-parameter sets of points by surfaces in $\mathbb{R}^{3,1}$ gives rise to general envelope surfaces of 2-parameter families of spheres, that have not been studied before in this generality. For this, a calculus was reinvented and applied, that classifies 2-planes in $\mathbb{R}^{3,1}$ according to their steepness.

Keywords: Minkowski space, numerical optimization, curve fitting, surface fitting, Laguerre geometry, cyclography

1. INTRODUCTION

Circles in \mathbb{R}^2 can be represented by points in Minkowski space $\mathbb{R}^{2,1}$. In order to find envelopes for the circles, we have to approximate the corresponding points by curves in $\mathbb{R}^{2,1}$, see Fig. 1. This would be a standard quadratic approximation problem if reality of envelopes would be ignored. But the envelopes are real if and only

if $\mathbf{c}(u)$ has no pseudo-Euclidean tangents, which is a non-convex quadratic constraint of the curve fitting.

1.1 Minkowski space

We only show some aspects of Minkowski space and Laguerre geometry; a more complete introduction to this classical topic can be found in

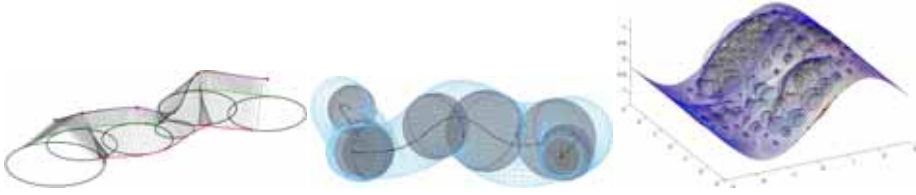


Figure 1: *left*: Given a set of circles in \mathbb{R}^2 , its envelope (curves in *red* and *green*) can be found as the inverse cyclographic image of an interpolating curve $\mathbf{c}(u)$ (*purple*) in Minkowski space $\mathbb{R}^{2,1}$ through points Q_k (*red* dots) representing the circles. The envelopes are real if and only if $\mathbf{c}(u)$ has no pseudo-Euclidean tangents, which is guaranteed through a constraint of the curve fitting. *middle*: One dimension higher, the input is a set of spheres (*gray*) for which we find the optimal envelope (*light blue* surface). *right*: Analogously, given a 2-parameter set of spheres (*gray*) what is the optimal surfaces enveloping them? These questions are answered through a non-convex constrained optimization which ensures reality of the envelopes.

[4].

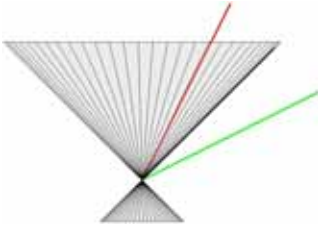


Figure 2: The scalar product $\langle \mathbf{a}, \mathbf{a} \rangle_L$ is negative/zero/positive if and only if the vector \mathbf{a} is *pseudo-Euclidean/isotropic/Euclidean* (colors *red / black / green*); the Lorentz cone is shown in *gray*.

Pseudo-Euclidean inner product. Minkowski space $\mathbb{R}^{2,1}$ is equipped with an indefinite inner product

$$\langle \mathbf{a}, \mathbf{b} \rangle_L = a_1 b_1 + a_2 b_2 - a_3 b_3 \quad (1)$$

and thus for a vector $\mathbf{a} \in \mathbb{R}^{2,1}$ $\langle \mathbf{a}, \mathbf{a} \rangle_L$ can be negative/zero/positive; the vector is then called *pseudo-Euclidean/isotropic/Euclidean*, see Fig. 2. The (convex) set $\{(x_1, x_2, x_3) \in \mathbb{R}^2 \times \mathbb{R} : x_3^2 \geq x_1^2 + x_2^2\}$ is called *Lorentz cone* Γ .

Quadratic constraint. Due to this geometric constraint, the tangents of these curves have to enclose an angle less than or equal to $\frac{\pi}{4}$ with the plane $x_3 = 0$. One can reformulate this constraint by saying that the hodograph (derivative curve) has to stay outside the Lorentz cone Γ ; it follows that this constraint is quadratic and non-convex.

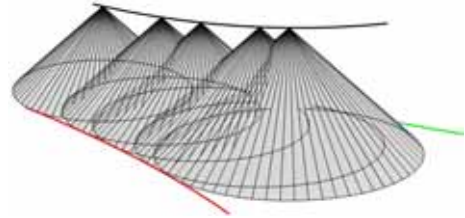


Figure 3: The cyclographic image of a curve $p(t)$ in Minkowski space $\mathbb{R}^{2,1}$ is the envelope of oriented circles. If the tangent direction $\dot{p}(t)$ encloses an angle with the plane $x_3 = 0$ that is more than/equal/less than $\frac{\pi}{4}$; we call the tangent line *pseudo-Euclidean/isotropic/Euclidean*. $\Rightarrow \zeta(p(t))$ is imaginary/one curve/two real curves.

1.2 Previous Work

The problem of finding envelopes of circles has been dealt with before, but not in this generality as we will see in a quick literature overview.

Pottmann and Peternell. [4] consider curves \mathbf{c} in $\mathbb{R}^{3,1}$ and interpret them as canal surfaces. If \mathbf{c} is a NURBS curve, its (*convex hull, variation diminishing*) properties translate into certain properties for the rational canal surfaces. Only curves with Euclidean tangents are considered, i.e. the question of reality of envelopes is not touched. The paper gives a full description of Dupin cyclides as images of intersections of three cones in $\mathbb{R}^{3,1}$. They also define *pe circular splines* and *pe biarcs*.

Slabaugh et al. have published [6] about envelopes of 1-parameter families of circles and [7] as a direct extension to 1-parameter families of spheres.

Their 2D method is to look at the two envelope curves separately and choose a point of contact S_i for every circle, which automatically gives a tangent direction T_i . Given the circles, these points of contact can be coded as an angle α_i . For two consecutive circles, a cubic Hermite interpolation is computed. This interpolation is then set up as an unconstrained quadratic optimization in the unknown α_i with minimization of arc length and curvature as objective function.

In 3D the points of contact are substituted by circles of contact, the rest extends straight forward.

The drawback of this method is that the curves are only C^1 and the knot vector of the piecewise cubic spline is determined by the rules of Hermite interpolation, thus the shape is not very flexible.

Kunkli and Hoffmann. [3] do a G^1 interpolation of circles via the circle of Apollonius: for three consecutive circles S_1, S_2, S_3 two other touching circles A_1, A_2 are constructed and the points of contact with A_1 are used for one envelope and the ones with A_2 for the other. Then, for two consecutive circles, tangents at the points of contact are computed and a Hermite interpolation is performed.

Their extension for spheres is to take the plane ε through the centers of three consecutive spheres

S_i , thus get three circles and once again apply Apollonius' method. Through two touching points a plane orthogonal to ε intersects S_i in a circle, which is used for another Hermite interpolation.

Such a Hermite interpolation is only a local construction and success is heavily dependent on the input.

2. ENVELOPES OF CIRCLES

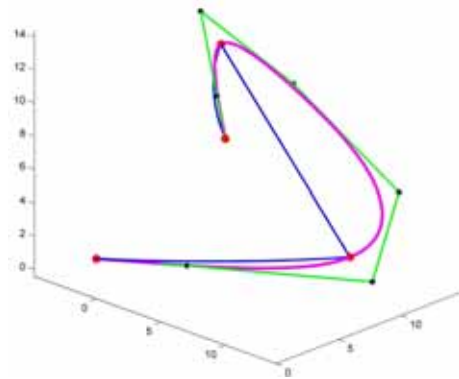


Figure 4: The original points in $\mathbb{R}^{2,1}$ in red and their connecting lines in blue - note that the blue line in the middle has an angle of $\frac{\pi}{4}$ with the xy -plane by construction. The final curve (purple) avoids steep tangents, as can be seen by the final control polygon in green.

Finding an envelope of an ordered set of circles in \mathbb{R}^2 can be translated by cyclography into the following task: Given a set of points Q_k in $\mathbb{R}^{2,1}$, find a curve $\mathbf{c}(u)$ that interpolates these. The inverse cyclographic mapping maps $\mathbf{c}(u)$ to the two branches of the circles' envelope, see Fig. 3.

A first look at the objective function. For given points Q_k , we want to

$$\text{minimize } \sum_{k=0}^{n-1} |Q_k - \mathbf{c}(u)|^2, \quad (2)$$

which means solving a least squares problem for unknown control points of a cubic B-spline curve

$\mathbf{c}(u)$. In fact, this approximation uses a footpoint, i.e. a point $\mathbf{c}(u_k)$ which changes in an iterative sub-routine, thus making (2) non-linear.

Note that the norm in eq. (2) is the Euclidean norm, because the Minkowski norm of eq. (1) can become zero even for isotropic vectors, which is impractical for this curve fitting.

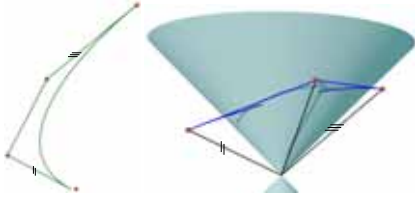


Figure 5: *left*: a cubic Bézier curve (green) and its control polygon, *right*: the hodograph (blue) is the curve of the first derivatives, which is a quadratic Bézier curve in the legs (black), i.e. the difference vectors. They originate at the origin of the Lorentz cone Γ (turquoise).

2.1 Introducing the constraint

The main challenge is that not every curve \mathbf{c} through points Q_k maps to a real envelope. As we have seen in sec. 1 this is only satisfied if

$$c_1'^2 + c_2'^2 \geq c_3'^2, \quad (3)$$

where $\frac{\partial \mathbf{c}}{\partial u} = \mathbf{c}' = (c_1', c_2', c_3') \in \mathbb{R}^{2,1}$ is called *hodograph*. Wherever (3) is strictly $>$, the envelopes do not coincide. The condition (3) is equivalent to saying

$\Leftrightarrow \langle \mathbf{c}', \mathbf{c}' \rangle_L \geq 0$ by using the Lorentz inner product of eq. (1), i.e. the derivative vectors are *Euclidean* everywhere.

\Leftrightarrow the angle between \mathbf{c}' and the xy -plane is $\leq \frac{\pi}{4}$,

\Leftrightarrow that the hodograph has to stay outside the Lorentz cone Γ , see Fig. 5.

This quadratic constraint is in fact non-convex as we will see shortly, thus standard optimization algorithms can not be applied. We will present a possibility to linearize it in sec. 2.2.

Quadratic objective in matrix notation. The objective function eq. (2) can be written as $(\mathbf{c}^T - \mathbf{q}^T) \cdot (\mathbf{c} - \mathbf{q}) \rightarrow \min$, thus

$$f(\mathbf{b}) = \mathbf{b}^T \cdot \mathbf{G} \cdot \mathbf{b} - 2 \cdot \mathbf{b}^T \cdot \mathbf{e} + \mathbf{q}^T \cdot \mathbf{q} \quad (4)$$

where $\mathbf{G} = \text{blkdiag}(\mathbf{N}^T \mathbf{N}, \mathbf{N}^T \mathbf{N}, \mathbf{N}^T \mathbf{N})$ is a symmetric block-diagonal matrix, $\mathbf{e} = \text{blkdiag}(\mathbf{N}^T, \mathbf{N}^T, \mathbf{N}^T) \cdot \mathbf{q}$ and $f(\mathbf{b})$

Solution via a linear system. To solve (4) iteratively, one writes $\mathbf{b}_{i+1} = \mathbf{b}_i + \Delta \mathbf{b}_i$ and arrives at an optimum for

$$\begin{aligned} \nabla_{\mathbf{b}} f &= \mathbf{G} \cdot (\mathbf{b}_i + \Delta \mathbf{b}_i) - \mathbf{e} = 0 \\ \Leftrightarrow \mathbf{G} \cdot \Delta \mathbf{b}_i &= \mathbf{e} - \mathbf{G} \cdot \mathbf{b}_i, \end{aligned}$$

which is a linear system in $\Delta \mathbf{b}_i$ for fixed parameters u_k . (??). We will assume to have found an optimum, whenever the absolute value of the change in error measure $\|\nabla f\|$ is smaller than some threshold.

2.2 Non-convex quadratic constraint

The curve $\mathbf{c} \in \mathbb{R}^{2,1}$ shall be constrained to have only Euclidean tangents, or equivalently, its cyclographic preimage shall be a real envelope to circles $\zeta^{-1}(Q_k)$. The derivative curve or *hodograph* is given as

$$[c_1' | c_2' | c_3'] = \mathbf{N}' \cdot [b_1 | b_2 | b_3],$$

with the same control points $\bar{\mathbf{b}} = [b_1 | b_2 | b_3]$ as the curve itself, and \mathbf{N}' denoting the collocation matrix of the derivatives of the basis functions. Then the constraint (3) can be reformulated as

$$\begin{aligned} b_1^T \mathbf{N}'^T \mathbf{N}' b_1 + b_2^T \mathbf{N}'^T \mathbf{N}' b_2 - b_3^T \mathbf{N}'^T \mathbf{N}' b_3 &\geq 0 \\ \Leftrightarrow \mathbf{b}^T \cdot \mathbf{A} \cdot \mathbf{b} &\geq 0, \end{aligned} \quad (5)$$

with $\mathbf{A} = \text{blkdiag}(\mathbf{N}'^T \mathbf{N}', \mathbf{N}'^T \mathbf{N}', -\mathbf{N}'^T \mathbf{N}')$. Note that the collocation matrix of the curve's derivative \mathbf{N}' does not need to be given at the same parameter values u_k , and the u_k do not need to change at every iteration.

Nevertheless, matrix \mathbf{A} is indefinite by construction, so methods like Quadratically constrained quadratic program (QCQP) cannot be applied. Note that since \mathbf{A} stems from evaluating basis functions at finitely many points, this condition is necessary, but not sufficient to satisfy the constraints.

2.3 Optimization procedure

Now that we have defined the matrices representing both objective function and the constraint in eq. (4) and (5) respectively, we can restate the original problem:

$$\text{minimize } \mathbf{b}^T \cdot \mathbf{G} \cdot \mathbf{b} - 2 \cdot \mathbf{b}^T \cdot \mathbf{e} + d \quad (6a)$$

$$\text{subject to } \mathbf{b}^T \cdot \mathbf{A} \cdot \mathbf{b} \geq 0, \quad (6b)$$

where \mathbf{G} is positive definite and \mathbf{A} is indefinite (both by construction).

Due to the indefiniteness of the constraint, (6) is a non-convex problem. If we take reparametrizations into account, which are necessary in order to include footpoints, the objective function is nonlinear.

Outline of the optimization procedure. To overcome these challenges, we will follow this procedure:

- (1) Start with a good initial position \mathbf{b}_0 for the curve's control polygon.
- (2) Choose footpoints $\mathbf{c}_i(u_k)$.
- (3) Compute a linearization \mathbf{D}_i of the quadratic constraint depending on the current control points \mathbf{b}_i .
- (4) Minimize the distance from \mathbf{c}_i to Q_k while staying feasible w.r.t \mathbf{D}_i and update the control points.

Repeat (2) - (4) until an optimum is reached.

We will make sure that the error introduced in (3) stays small by iteratively adapting the linearization in (2) and (3). Therefore, step (4) turns the non-convex, quadratically constraint problem

(6) into a Quadratic Program. This important linearization will be carried out in the next section 2.4.

2.4 Linearizing the quadratic constraint

In this section we will present a local linearization and proceed in three steps:

- Show in sec. 2.4 that bounding the samples of the hodograph away from the Lorentz cone can locally be accomplished through Γ 's tangent planes.
- Present a routine that computes the corresponding tangent plane for each sample in sec. 2.4 with the help of a projection
- Give a matrix formulation of the linearized constraint that only depends on the control points of the last iterate in sec. 2.4.

Local linearization of Γ . As was said in the introduction, the quadratic constraint is equivalent to the hodograph \mathbf{c}' staying outside the Lorentz cone Γ .

Note that the space $\mathbb{R}^{2,1} \setminus \Gamma$ is not a vector space; a linear combination of two Euclidean vectors is not necessarily Euclidean. So rather than giving a condition on the *control polygon* of the hodograph, we look at *sampling points* $\mathbf{c}'(u_k)$.

The linearization for condition " $\mathbf{c}'(u_k)$ must stay outside Γ " will be " $\mathbf{c}'(u_k)$ must stay on the positive side of a tangent plane of Γ ". The former can be written as $c_1'^2 + c_2'^2 - c_3'^2 \geq 0$, the latter as $\bar{p}_1 c_1' + \bar{p}_2 c_2' - \bar{p}_3 c_3' \geq 0$ for $\bar{\mathbf{p}} = (\bar{p}_1, \bar{p}_2, \bar{p}_3)$ on the cone.

This connection comes naturally when looking at the equation of the border of the Lorentz cone $\partial\Gamma$, which is given by

$$\partial\Gamma : x^2 + y^2 - z^2 = 0 \quad (7)$$

thus the normal vector of its (isotropic) tangential plane ε is

$$\nabla(x^2 + y^2 - z^2) = (2x, 2y, -2z) = (x, y, -z).$$

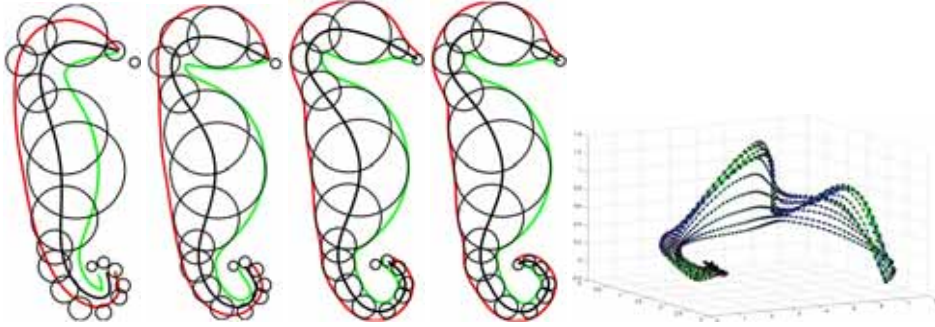


Figure 6: An example of envelopes using the Interior Point algorithm at different stages, *top from left to right*: fifth iteration after initialization with a flat curve, after ten iterations, after 15 iterations, final result after 30 iterations. *right*: curves in $\mathbb{R}^{2,1}$ for the first 15 iterations.

Therefore, given a point $\bar{\mathbf{p}}$ on the cone it is contained in the plane

$$\varepsilon: \bar{p}_1x + \bar{p}_2y - \bar{p}_3z = 0 \quad (8)$$

For this tangent plane, the normal is pointing away from Γ and thus the linear form $\langle \bar{\mathbf{p}}, \mathbf{x} \rangle_L$ derived from eq. (8) is positive for \mathbf{x} close to $\bar{\mathbf{p}}$ and outside Γ . This linear form is also a good local approximation of the quadratic form $\langle \mathbf{x}, \mathbf{x} \rangle_L$ for a suitable $\bar{\mathbf{p}}$.

Projection orthogonal to Γ . We have seen in the last section that a local linearization of Γ depends on the choice of a suitable point $\bar{\mathbf{p}} = (\bar{p}_1, \bar{p}_2, \bar{p}_3)$ on the cone for a point $\mathbf{p} = (p_1, p_2, p_3)$, such that $\langle \bar{\mathbf{p}}, \mathbf{p} \rangle_L$ is positive/negative/zero for a Euclidean/pseudo-Euclidean/isotropic point.

Keep in mind that these "points" are actually *derivative vectors* and hence belong to the tangent space of $\mathbb{R}^{2,1}$. A suitable choice is

$$\bar{\mathbf{p}} = \mathbf{p} - \lambda \cdot \frac{\mathbf{n}_\varepsilon}{\|\mathbf{n}_\varepsilon\|}, \quad \lambda = \frac{p_1^2 + p_2^2 \mp p_3 \sqrt{p_1^2 + p_2^2}}{\sqrt{2(p_1^2 + p_2^2)}} \quad (9)$$

for the smaller of the two λ 's.

2.5 Matrix formulation of the linearization

We are now ready to write the linearization of the quadratic constraint (5) in a matrix formulation that can be used in an optimization algorithm.

We have seen in sec. 2.4 that for a derivative vector $\mathbf{d}_k = \mathbf{c}'(u_k) = \mathbf{N}'(u_k) \cdot \bar{\mathbf{b}}$ the quadratic constraint $\langle \mathbf{d}_k, \mathbf{d}_k \rangle_L \geq 0$ can be linearized as

$$\begin{aligned} \langle \bar{\mathbf{d}}_k, \mathbf{d}_k \rangle_L &= [\bar{d}_{k1}, \bar{d}_{k2}, \bar{d}_{k3}] \text{diag}(1, 1, -1) \begin{bmatrix} d_{k1} \\ d_{k2} \\ d_{k3} \end{bmatrix} \\ &= [\bar{d}_{k1} \cdot \mathbf{N}'(u_k), \bar{d}_{k2} \cdot \mathbf{N}'(u_k), -\bar{d}_{k3} \cdot \mathbf{N}'(u_k)] \cdot \mathbf{b} \geq 0, \end{aligned}$$

with $\bar{\mathbf{d}}_k$ the projection of \mathbf{d}_k onto Γ as described in eq. (9), $\mathbf{N}'(u_k)$ one row of the collocation matrix of the hodograph \mathbf{c}' and \mathbf{b} the $3m \times 1$ -vector of the m control points of \mathbf{c} .

If we want to carry out this operation for all \mathbf{d}_k with a single matrix multiplication, we have to write the $\bar{\mathbf{d}}_k$ coordinate-wise in diagonal matrices and multiply these with the $n \times m$ -matrix \mathbf{N}' :

$$\mathbf{D} := [\text{diag}(\bar{d}_{11} \dots \bar{d}_{1n})\mathbf{N}', \text{diag}(\bar{d}_{21}, \dots, \bar{d}_{2n})\mathbf{N}', \\ - \text{diag}(\bar{d}_{31}, \dots, \bar{d}_{3n})\mathbf{N}']$$

and get the matrix (=linear form) linearizing the quadratic form of the constraint

$$\mathbf{D} \cdot \mathbf{b} = [\langle \bar{\mathbf{d}}_1, \mathbf{d}_1 \rangle_L, \dots, \langle \bar{\mathbf{d}}_n, \mathbf{d}_n \rangle_L]^T. \quad (10)$$

Remember that an entry $\langle \bar{\mathbf{d}}_k, \mathbf{d}_k \rangle_L$ in (10) is actually the signed distance for a point $\mathbf{c}'(u_k)$ to a certain (closest) tangent plane of Γ .

Also keep in mind that even though \mathbf{N}' might remain unchanged throughout the iteration, \mathbf{D}_i depends on the control points \mathbf{b}_{i-1} of the last iterate, because $\mathbf{c}'_i = \mathbf{N}' \cdot \mathbf{b}_{i-1}$.

2.6 Initial position

We would like to have a much bigger number m of control points \mathbf{b} than the number n of points $Q_k \in \mathbb{R}^{2,1}$ representing the circles in order to have more flexibility in optimization. Trying to solve this directly would lead to a rank-deficient linear system whose solution \mathbf{b} can, of course, not be trusted.

To circumvent this problem, we introduce an auxiliary cubic B-spline curve \mathbf{c}_{aux} , interpolating the input Q_k with a minimal knot vector and ignoring the steepness constraint, see Fig. 7.

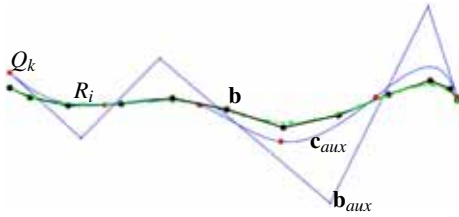


Figure 7: Procedure to find the initial position of the control polygon: Compute auxiliary curve \mathbf{c}_{aux} (blue curve) with control polygon \mathbf{b}_{aux} (blue lines) through input points Q_k (red dots). Move this control polygon toward the mean of the z -coordinates of the Q_k until no steep tangents appear, then sample points R_i (green dots). A finer control polygon \mathbf{b} (black) for approximating those is found via a linear system.

We take $m_1 > m + 4$ equally spaced auxiliary parameter values v_i , a knot vector of length $m + 4$ and compute thus the collocation matrix \mathbf{N}_{aux} of rank m .

The z -coordinates of the control points \mathbf{b}_{aux} of \mathbf{c}_{aux} are then moved half the way to the

mean value of the z -coordinates of the Q_k , called $\text{mean}(Q_{k,z})$. Should \mathbf{c}_{aux} still have pseudo-Euclidean tangents, repeat this scaling; sample the final Euclidean curve $R_i = \mathbf{c}_{aux}(v_i)$.

The initial control polygon \mathbf{b} for the optimization procedure is the least squares fit of the over-determined system $\mathbf{N}_{aux} \cdot \mathbf{b} = R_i$. It is Euclidean by construction and the number of control points m is a user-specified number.

Note that due to the repeated scaling of the z -coordinates of \mathbf{c}_{aux} , the B-Spline defined by \mathbf{b} can be quite far from the Q_k (only in z -direction); in the worst case, it is approximating the top view projection of the Q_k in the plane $z = \text{mean}(Q_{k,z})$.

3. ENVELOPES OF SPHERES

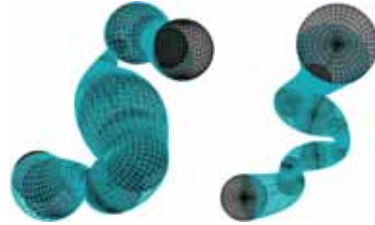


Figure 8: Two examples of envelopes of spheres using the Interior Point algorithm generalized to $\mathbb{R}^{3,1}$, e.g. canal surfaces. The original spheres are shown in black, the resulting envelope in turquoise. The curve approximation and the steepness constraint carry over directly from $\mathbb{R}^{2,1}$ to $\mathbb{R}^{3,1}$.

The direct extension of the theory of section 2 to the 4-dimensional $\mathbb{R}^{3,1}$ would be curves in $\mathbb{R}^{3,1}$, which represent envelopes of 1-parameter families of spheres, i.e. canal surfaces. Basically, the theory stays the same as in section 2; we refer the interested reader to [1] to see which formulas have to be adapted to the higher dimension.

The main contribution of the section lies with surfaces in $\mathbb{R}^{3,1}$, whose inverse cyclographic image are envelopes of 2-parameter families of spheres. The constraint of having real envelopes means that the tangent planes of these surfaces

have to stay Euclidean. In fact, we are talking of 2-planes in 4-space and section 3.1 introduces a calculus that allows to classify such planes through a bilinear form.

3.1 Wedge product in 4-space

We want to define the wedge product $\mathbf{x} \wedge \mathbf{y}$ of two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^4$ (or $\mathbb{R}^{3,1}$; the differences between those two exterior algebras solely depend on the choice of basis), which is a vector in the $\binom{4}{2} = 6$ -dimensional vector space $\bigwedge^2 \mathbb{R}^4$.

Let $(\mathbf{e}_0, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ be a basis of \mathbb{R}^4 , then a basis of $\bigwedge^2 \mathbb{R}^4$ is given as $(\mathbf{e}_0 \wedge \mathbf{e}_1, \mathbf{e}_0 \wedge \mathbf{e}_2, \mathbf{e}_0 \wedge \mathbf{e}_3, \mathbf{e}_1 \wedge \mathbf{e}_2, \mathbf{e}_1 \wedge \mathbf{e}_3, \mathbf{e}_2 \wedge \mathbf{e}_3)$. This linear space is the set of all 2-spaces (=planes through the origin) in 4-space.

Thus for two vectors $\mathbf{x} = (x_0, x_1, x_2, x_3)$, $\mathbf{y} = (y_0, y_1, y_2, y_3)$ the coefficients of the wedge product $\mathbf{x} \wedge \mathbf{y}$ are given by the numbers $l_{ij} = x_i y_j - x_j y_i$ for $i, j = 0, \dots, 3, i \neq j$, and the coordinates of the plane through the origin spanned by $\mathbf{x} \wedge \mathbf{y} \in \bigwedge^2 \mathbb{R}^4$ are $L = (l_{01}, l_{02}, l_{03}, l_{23}, l_{31}, l_{12})$. Note that a vector L satisfies the Plücker identity

$$\Omega_q(L) = l_{01}l_{23} + l_{02}l_{31} + l_{03}l_{12} = 0, \quad (11)$$

if and only if it represents a 2-plane in \mathbb{R}^4 ; see [5] and the use of this formalism in *line geometry*. We heavily rely on the fact that (projective) lines in projective three-space \mathcal{P}^3 are isomorphic to 2-planes through the origin in \mathbb{R}^4 .

The bilinear form $\Phi_{3,1}$. We want to relate inner product and wedge product in $\bigwedge^2 \mathbb{R}^4$.

Lemma 1. *The Plücker coordinates of a plane $\mathbf{u} \wedge \mathbf{v}$ be l_{ij} for $\mathbf{u}, \mathbf{v} \in \mathbb{R}^{3,1}$ and those of $\mathbf{x} \wedge \mathbf{y}$ be m_{ij} . The following holds*

$$\begin{aligned} \det \begin{pmatrix} \langle \mathbf{u}, \mathbf{x} \rangle_L & \langle \mathbf{v}, \mathbf{x} \rangle_L \\ \langle \mathbf{u}, \mathbf{y} \rangle_L & \langle \mathbf{v}, \mathbf{y} \rangle_L \end{pmatrix} &= \\ &= l_{01}m_{01} + l_{02}m_{02} + l_{12}m_{12} \\ &- l_{03}m_{03} - l_{13}m_{13} - l_{23}m_{23} \end{aligned}$$

We define a bilinear form $\Phi_{3,1} : \bigwedge^2 \mathbb{R}^{3,1} \times$

$\bigwedge^2 \mathbb{R}^{3,1} \rightarrow \mathbb{R}$ by setting

$$\begin{aligned} \Phi_{3,1}(\mathbf{u} \wedge \mathbf{v}, \mathbf{x} \wedge \mathbf{y}) &:= l_{01}m_{01} + l_{02}m_{02} + l_{12}m_{12} \\ &- l_{03}m_{03} - l_{13}m_{13} - l_{23}m_{23}, \end{aligned}$$

where l_{ij} are the Plücker coordinates of a plane $\mathbf{u} \wedge \mathbf{v}$ and m_{ij} those of $\mathbf{x} \wedge \mathbf{y}$ as in Lemma 1.

Let $\mathbf{u}, \mathbf{v} \in \mathbb{R}^{3,1}$ be two vectors. Then we get the important relation

$$\Phi_{3,1}(\mathbf{u} \wedge \mathbf{v}, \mathbf{u} \wedge \mathbf{v}) = \langle \mathbf{u}, \mathbf{u} \rangle_L \cdot \langle \mathbf{v}, \mathbf{v} \rangle_L - \langle \mathbf{u}, \mathbf{v} \rangle_L^2.$$

This is analogous to the lower dimensional case, which serves as a classification of 1-planes in $\mathbb{R}^{2,1}$, see [1].

3.2 Classification of 2-planes in Minkowski space

In this section we present a classification of 2-planes in $\mathbb{R}^{3,1}$ that fits consistently to the lower dimensional case, and extend a proposition on the classification of 1-planes in Minkowski space $\mathbb{R}^{2,1}$.

Lemma 2. *The bilinear form $\Phi_{3,1}$ allows for classification of 2-planes in $\mathbb{R}^{3,1}$:*

1. $\mathbf{u} \wedge \mathbf{v}$ is an Euclidean plane
 $\Leftrightarrow \Phi_{3,1}(\mathbf{u} \wedge \mathbf{v}, \mathbf{u} \wedge \mathbf{v}) > 0$
 $\Leftrightarrow l_{01}^2 + l_{02}^2 + l_{12}^2 > l_{03}^2 + l_{13}^2 + l_{23}^2$
2. $\mathbf{u} \wedge \mathbf{v}$ is an isotropic plane
 $\Leftrightarrow \Phi_{3,1}(\mathbf{u} \wedge \mathbf{v}, \mathbf{u} \wedge \mathbf{v}) = 0$
 $\Leftrightarrow l_{01}^2 + l_{02}^2 + l_{12}^2 = l_{03}^2 + l_{13}^2 + l_{23}^2$
3. $\mathbf{u} \wedge \mathbf{v}$ is a pseudo-Euclidean plane
 $\Leftrightarrow \Phi_{3,1}(\mathbf{u} \wedge \mathbf{v}, \mathbf{u} \wedge \mathbf{v}) < 0$
 $\Leftrightarrow l_{01}^2 + l_{02}^2 + l_{12}^2 < l_{03}^2 + l_{13}^2 + l_{23}^2$

Note that this lemma first appeared in [2], and was reinvented in [1], where an alternative proof was found.

We will use this classification of 2-planes in sec. 3.3 to ensure that a tangent plane to a surface in $\mathbb{R}^{3,1}$ is Euclidean in an optimization routine.

3.3 Surface optimization

We want to turn our attention to surfaces in $\mathbb{R}^{3,1}$, e.g. the optimization of envelopes for 2-parameter families of spheres.

The optimization problem thus becomes

$$\text{minimize } \sum_{i=1}^k |Q_i - \mathbf{f}(u, v)|^2 \quad (12a)$$

$$\text{subject to } \zeta^{-1}(\mathbf{f}(u, v)) \text{ is real,} \quad (12b)$$

and the distance in the objective function (12a) is the *Euclidean* distance, rather than the distance based on the *Minkowski norm* of eq. (1). In the latter case we would run into all sorts of problems with isotropic directions, which we want to avoid altogether.

The constraint (12b) means that the inverse cyclographic image of the surface parametrization in \mathbb{R}^3 stays real, or, equivalently, that the tangent planes of $\mathbf{f}(u, v) \subset \mathbb{R}^{3,1}$ stay Euclidean.

The formulation of this constraint is more involved than in the curve case and we will explain it in several steps in sec. 3.4.

Computation and parametrization. We have introduced a formalism in sec. 3.2 that allows us to check whether a 2-plane in $\mathbb{R}^{3,1}$ is pseudo-Euclidean, isotropic or Euclidean; here we want to apply it to tangent planes of a surface. We must therefore choose a parametrization $\mathbf{f}(u, v)$ and for practical reasons we restrict our attention to tensor product B-spline surfaces.

Handling the constraint - 3 levels of rigidity. For three different classes of surfaces we present three levels of rigidity for the steepness constraint:

1. The most general class of surfaces we will look at are *bicubic tensor product B-spline surfaces* $\mathbf{f}(u, v)$, ie. the parameter lines in u - and v -direction are all cubic B-spline curves. On these, we use a sampling of the surface and can only guarantee that the tangent planes are Euclidean at each sampling

point, which means that the envelope of a 2-parameter family of spheres $\zeta^{-1}(\mathbf{f}(u, v))$ is real at a discrete number of points. This general approach, which is the foundation of all three, will be explained in sec. 3.4.

2. If we consider tensor product B-splines surfaces of bidegree (3, 1), ie. strips of *ruled surfaces*, one direction of parameter lines, say u -lines, will be cubic B-splines and the other direction linear B-splines, ie. polylines. We will sample the cubic u -parameter lines and given that they are Euclidean (on a sampling), the tangent planes along the rulings (= the v -parameter lines) can be guaranteed to be Euclidean - without the use of a sampling.
3. If both directions are linear, the single patches of the tensor product surface are *hyperbolic paraboloids* in $\mathbb{R}^{3,1}$. For this class of surfaces, we can even guarantee Euclidean surfaces without the need of a sampling.

For brevity, we will only described the most general class here and refer to [1] for the other two.

3.4 Constraints for surface optimization

Analogously to sec. 2.4, we have to give a linearized version of the quadratic constraint (12b), which in the surface case is actually slightly more involved.

The constraint should be, that for a surface $\mathbf{f}(u, v)$ its tangent planes stay Euclidean everywhere. We will sample the parameter lines to check the steepness of the partial derivatives.

We describe this constraint in the following steps:

1. We ensure Euclidean tangents in the direction of the parameter lines, i.e. the constraint shall be fulfilled for a sampling of the partial derivatives $\mathbf{f}_u(u_i, v_j)$ and $\mathbf{f}_v(u_i, v_j)$.

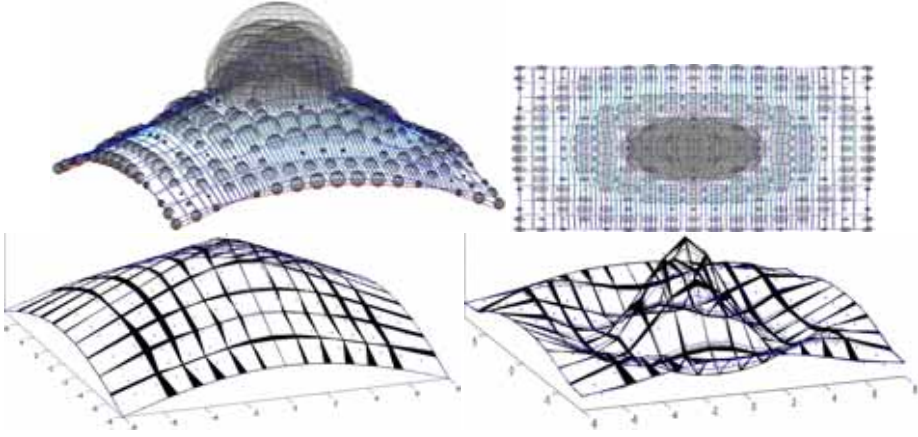


Figure 9: The surface $f(u, v) = (u, v, 3 - \frac{r}{2}, 3 \frac{\sin(r)}{r})$: *top row left*: an axonometric view of the hyperplanar section $x_3 = 0$, *right*: an axonometric view of the hyperplanar section $x_2 = 0$; surface evaluated on a grid, points with pseudo-Euclidean (mixed) partial derivatives in *red*; *second row*: the cyclographic preimage of an initial value for a bicubic B-spline surface approximation; *last row left*: values in the hyperplanar section $x_3 = 0$, *right* in hyperplanar section $x_2 = 0$ from start (*blue*) and intermediates to finish (*black*).

2. If both $\mathbf{f}_u(u_i, v_j)$ and $\mathbf{f}_v(u_i, v_j)$ are Euclidean, we test if the tangent plane at $\mathbf{f}(u_i, v_j)$ is Euclidean by computing $\Phi_{3,1}(\mathbf{f}_u \wedge \mathbf{f}_v, \mathbf{f}_u \wedge \mathbf{f}_v)(u_i, v_j)$. We give a matrix formulation for constraining the tangent plane of a surface in $\mathbb{R}^{3,1}$ to be Euclidean.

For the special cases of strips of ruled surfaces and patches of hyperbolic paraboloids, further steps are necessary, see [1].

The constraint on parameter lines. We assume $\mathbf{f}(u, v)$ is parametrized as a cubic tensor product B-spline surface, hence the parameter lines in u and v -direction are cubic B-spline curves.

Therefore, constraining them to have only Euclidean tangents is exactly constraint (12b), i.e. the constraint for a curve in $\mathbb{R}^{3,1}$ to have only Euclidean tangents.

Then, for k sample points on the surface $\mathbf{f}(u, v)$, the $k \times 4$ -matrix of partial derivatives w.r.t. u is

given as $\mathbf{f}_u = \mathbf{N}_{du} \cdot \mathbf{b}$ and we define $\tilde{\mathbf{f}}_u$ to be the projection - again a $k \times 4$ -matrix - of \mathbf{f}_u onto Γ as defined in sec. 2.4.

Then $\mathbf{D}_u \cdot [\mathbf{b}_x^T | \mathbf{b}_y^T | \mathbf{b}_z^T | \mathbf{b}_w^T]^T = \langle \mathbf{f}_u, \tilde{\mathbf{f}}_u \rangle_L$ is a vector, whose k th entry is less than zero if and only if $\mathbf{f}_u(u_k, v_k)$ is pseudo-Euclidean.

The matrix \mathbf{D}_u thus linearizes the quadratic constraint that the angle between the partial derivatives w.r.t u , $\mathbf{f}_u(u_k, v_k)$, and the hyperplane $x_4 = 0$ is less than $\frac{\pi}{4}$ for a sampling of all parameter lines in u -direction.

The constraint for a discrete sampling of the surface. Knowing that the partial derivatives $\mathbf{f}_u(u_i, v_j)$ and $\mathbf{f}_v(u_i, v_j)$ are Euclidean does not necessarily mean that the tangent plane at $\mathbf{f}(u_i, v_j)$ is Euclidean.

We thus employ the calculus summarized in Lemma 2: Compute $\Phi_{3,1}(\mathbf{f}_u \wedge \mathbf{f}_v, \mathbf{f}_u \wedge \mathbf{f}_v)(u_i, v_j)$; if it is ≤ 0 , the tangent plane at $\mathbf{f}_v(u_i, v_j)$ is isotropic or pseudo-Euclidean, even though the

dual variables of the optimization are positive by assumption. We hence change their sign and thus activate this constraint.

Algorithm for constraining the tangent planes of a surface. Let us summarize the ideas of this section as an algorithm:

1. For a sampling of a B-spline surface $\mathbf{f}(u_i, v_j)$, compute the matrices \mathbf{D}_u and \mathbf{D}_v of the linearized constraint.
2. For index pairs (\bar{i}, \bar{j}) for which the slack variables > 0 , compute the bilinear form $\Phi_{3,1}(\mathbf{f}_u \wedge \mathbf{f}_v, \mathbf{f}_u \wedge \mathbf{f}_v)(u_{\bar{i}}, v_{\bar{j}})$.
3. For those (\bar{i}, \bar{j}) for which $\Phi_{3,1} \leq 0$, change the sign of the slack variable to be minus.

With the simple trick in 3., which is at the heart of the primal-dual interior point algorithm used to solve these problems, we have activated the constraint with the help of $\Phi_{3,1}$, even though both partial derivatives are Euclidean. This way, isotropic or pseudo-Euclidean tangent planes enter the optimization.

4. CONCLUSIONS

This paper summarizes the author's PhD thesis [1], which finds envelopes for circles in the plane and for 1- and 2-parameter families of spheres in 3-space via a nonlinear quadratic optimization in Minkowski space. Methods for handling a non-convex constraint are employed and results are presented through examples.

ACKNOWLEDGMENTS

I want to thank Martin Peternell who was the supervisor of my PhD thesis [1], of which this paper is a condensed version.

REFERENCES

- [1] B. Blaschitz. *Geometric Optimization in Minkowski Space*. Ph.D. thesis, Vienna University of Technology, 2014.
- [2] J. Kosinka and B. Jüttler. MOS surfaces: Medial surface transforms with rational domain

boundaries. In *Mathematics of Surfaces XII*, pages 245–262. Springer, 2007.

- [3] R. Kunkli and M. Hoffmann. Skinning of circles and spheres. *Computer Aided Geometric Design*, 2010.
- [4] H. Pottmann and M. Peternell. Applications of Laguerre geometry in CAGD. *Computer Aided Geometric Design*, 15(2): 165–186, 1998.
- [5] H. Pottmann and J. Wallner. *Computational Line Geometry*. Springer, 2001.
- [6] G. Slabaugh, G. Unal, T. Fang, J. Rossignac, and B. Whited. Variational skinning of an ordered set of discrete 2D balls. In *Proceedings of the 5th international conference on Advances in geometric modeling and processing*, pages 450–461. Springer-Verlag, 2008.
- [7] G. Slabaugh, B. Whited, J. Rossignac, T. Fang, and G. Unal. 3D ball skinning using PDEs for generation of smooth tubular surfaces. *Computer-Aided Design*, 42(1): 18–26, 2010.

ABOUT THE AUTHOR

Bernhard Blaschitz wrote his PhD thesis under the supervision of Martin Peternell at the Institute of Discrete Mathematics and Geometry at Technical University of Vienna, where he was a teaching assistant. He now works for Waagner-Biro's Advanced Geometry Engineering team.

GEOMETRIC PERSPECTIVE - CYLINDRICAL PERSPECTIVE

Flamur Doli

Faculty of Civil Engineering and Architecture, University of Prishtina, Republic of Kosova

ABSTRACT: This study is about the basics of cylindrical perspective construction that represents the central projection in semicylindrical surfaces. In this perspective, the projection surface is semicylindrical, while the basic plane of projection is in frontal position and passes through the axis of projection surface, where the center of projection- the eye point (O) is also found.

In contrast to linear perspective, in cylindrical perspective the construction of the perspective figure (perspective) is not carried in the plane of figure, but in a semicylindrical projection surface. In the linear perspective, the plane of figure is the same with the basic plane of projection, in which the perspective of figure is directly constructed, while in cylindrical perspective, the object perspective figure already constructed in a semicylindrical projection surface is projected on the basic plane of projection.

Cylindrical perspective is in a higher level compared to the linear one, as the cylindrical perspective of the line is not still a line, but a spatial curve on a projection surface.

Keywords: Cylindrical perspective, semicylindrical surface, central projection, basic plane of projection, spatial curve.

1. CYLINDRICAL PERSPECTIVE ELEMENTS

Cylindrical perspective or the central projection on a cylindrical surface is constructed through methods of descriptive geometry. Elements of the cylindrical perspective are: object; projection surface in which the central projection is constructed or figure perspective of an object; center of projection- eye point O, through which all the projection rays pass-viewing rays; the basic plane of projection in which the figure perspective of an object is projected that is already constructed in the projection surface. Basic plane of projection has a frontal position and passes through the eye point O. Hence, this plane is named the eye plane and is considered to coincide with the projection sheet. In this context, the space

ahead the eye plane presents the real figure space or the “view” space, while the space beyond this plane is named the geometric space. Beside these elements, the horizon plane H and the projection axis are also found, both these elements pass through the eye point O. Horizon plane is otherwise known as the eye point height plane or the eye horizontal. The basic plane of projection or the horizontal is found at a distance (h) under and parallel to the horizon plane. Intersection of this plane with the basic plane of projection determines the basic line x, whereas the line where the horizon and basic plane intersect determines the horizon line h.

2. THE CYLINDRICAL PERSPECTIVE OF A POINT

The point where the viewing ray PO and the semicylindrical projection surface intersect de-

noted as P_c determines the perspective figure or the cylindrical perspective of point P as seen on fig.1. This intersecting point is determined through vertical plane OPP_cP_c' that intersects

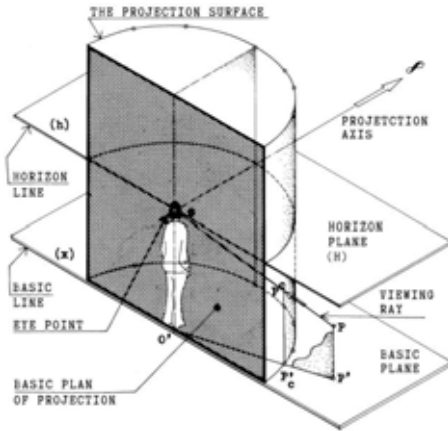


Figure 1: The cylindrical perspective elements, the cylindrical perspective of a point.

with the projection surface according to the vertical P_cP_c' : the point where the vertical P_cP_c' and viewing ray PO intersect presents the cylindrical perspective P_c of point P . The perspective figure P_c of point P is also the cylindrical perspective of all other points belonging to the viewing ray PO . Therefore any point, line or object must have a specific position in space. Point A on fig.2.a) is given with a certain position defined in space, determined by: the first orthogonal projection of point A' on the basic (horizontal) plane and the second orthogonal projection A'' on the basic (frontal) plane of projection. Such planes of projection can be considered as the first-horizontal plane of projection and the second-frontal plane of projection. Thus, the vertical distance AA' of point A from the basic plane is labeled as the height of point, whereas the horizontal distance AA'' of point A from the basic plane of projection is labeled as the point distance.

Figure 2 presents both orthogonal projections of the viewing ray AO on a single projection plane- created after the rotation of the basic plane (BP) around the basic line x and the bottom part of basic plane V with which it complies. This means that from the spatial system of the two orthogonal planes of projection we have proceeded in a single projection plane in a frontal position that coincides with our projection sheet. Perspective figure- the cylindrical perspective of any point, line, geometric figure or any other object as we will see later, is always constructed on this projection plane and above the basic line x . The construction of the cylindrical perspective of the upper mentioned objects follows as below:

2.1.

Determination of the cylindrical perspective of point A consists in determining the point where viewing ray AO intersects with the semicylindrical projection surface, fig.2.a). The axis OO' of this surface in which eye point O is found, lies on the basic projection plane V . In the spatial scheme, fig.2.a) is seen that the first projection $A'c'$ of the intersecting point $A'c'$ of the viewing ray AO with the projection surface, lies on the first (basic) projection of the semicylindrical projection surface. Accordingly, the first projection $A'c'$ of the cylindrical perspective of point A , [see fig.2.b)], lies on the intersection of the first projection of the viewing ray $A'O'$ and the first projection of the semicylindrical projection surface. While, the cylindrical perspective A_c of point A is constructed by drawing a vertical starting from $A'c'$ until it intersects with the second projection of the viewing ray $A''O''$. The cylindrical perspective of any point is determined by first obtaining its first projection and then drawing a vertical from this projection until it intersects with the second projection of the viewing ray, where the required cylindrical perspective of a point is found. Hence, in constructing cylindrical perspectives of points their first projections have to be determined first. Again,

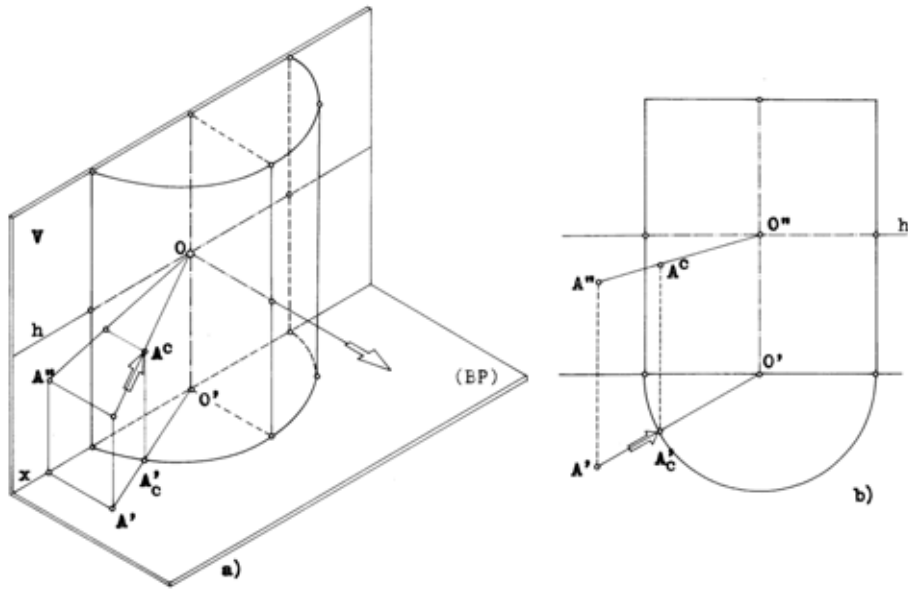


Figure 2: Determination of the cylindrical perspective of a point

the perspective of a point lies on the second projection of the viewing ray or on the line drawn from the second projection O'' of eye point O and the second projection of the given point.

3. CYLINDRICAL PERSPECTIVE OF THE LINE AND ITS INFINITE POINT

Cylindrical perspective of a line is determined by the geometric position of cylindrical perspectives of all its points. This means that in order to determine the cylindrical perspective of a line, the cylindrical perspective of each of its constituent point has to be determined first and by joining these perspectives the cylindrical perspective of a line is constructed.

3.1. Cylindrical perspective of lines with a specific position

As it is seen on figures 4.a) and 4.b) in order to find the cylindrical perspective of a given line dn all the perspectives of its constituent points have to be determined first through their viewing rays. Thus, the cylindrical perspectives of points 1, 2, 3 are found at the intersection of the corresponding view rays with the cylindrical projection surface of each point and by joining the cylindrical perspectives $1c$, $2c$, $3c$, the cylindrical perspective dnc of line dn is constructed.

3.1.1. The infinite point of a line

Viewing rays that unite points of a line with the eye point O are characterized by a ray that unites the eye point with a point of a line that is infinitely distanced from the basic projection plane.

As it is seen on fig.4.a) this ray shares a parallel

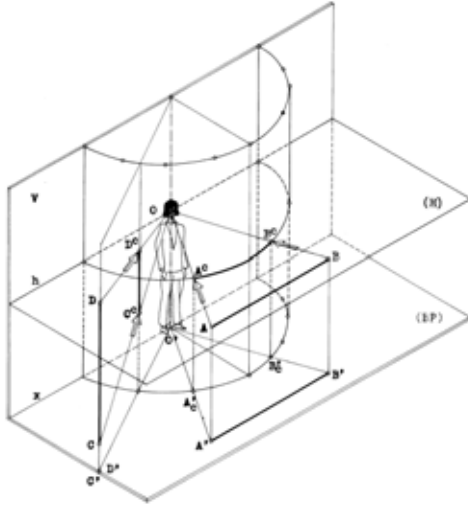


Figure 3: Determination of a segment perspective (AB and BC) on the projection surface.

position with a given line and in perspective it is called a parallel ray, denoted by s_d . The infinite point of a line is found at the intersection of this parallel ray s_d with the projection surface denoted by I_{dn} . From this, it is seen that the cylindrical perspective of a line is also determined through its infinite point. In perspective, two or more lines lying parallel in space share a common infinite point. [see fig. 5.a) and 5.b]

3.1.2. The plane of viewing rays

There is another method to determine the cylindrical perspective of line which integrates a new notion: the plane of rays or the viewing plane. Viewing rays that unite the constituent points of line dn with the eye point O create the viewing plane. In fact, the viewing plane is represented by the plane that passes through the given line dn and the eye point O. The cylindrical perspective of a line dnc is found where the plane of rays intersects with the

semicylindrical projection surface. It is clear that this cylindrical perspective of a line projected in one projection plane, respectively in the basic projection plane presents as a line [fig.4.b) line dnc]. The cylindrical perspective of line do with a specific position- that coincides with the projection axis- is a point. Therefore, the cylindrical perspective of the projection axis is also a point.

3.1.3.

Figure 5 presents a method of determining the cylindrical perspective for two lines with a specific position- a horizontal position that is perpendicular to the basic projection plane V. Viewing plane of line $d1$ is created by the viewing rays uniting its points with eye point O. This plane has exactly the same position with the horizon plane (H). So, the cylindrical perspective not only of line $d1$ or $d2$, but of all points, lines and other geometric figures that lie in the horizon plane (H) are found on the horizon line h [see fig.5.b]. with a specific position.

3.1.4.

In figure 6, lines $dh1$ and $dh2$ are in a horizontal position and in the mean time are parallel with the basic projection plane. The cylindrical perspective of these lines is determined by a line that is derived from the intersection of the semicylindrical projection surface with the viewing planes that pass through eye point O and lines $dh1$ and $dh2$. Infinite points, Ia and Ib of these lines are found at the intersection of the parallel ray with the projection surface, more specifically with the generatrices of this surface that lie on the basic projection surface V. So, the parallel ray- viewing ray that unites eye point O, with points M and N of lines $dh1$ and $dh2$ which lie infinitely on both opposite directions of a line, lies on the basic projection plane V and takes a parallel position with the basic line x . Infinite points of these lines Ia and Ib are found where parallel rays intersect with the projection surface, exactly at the vertical

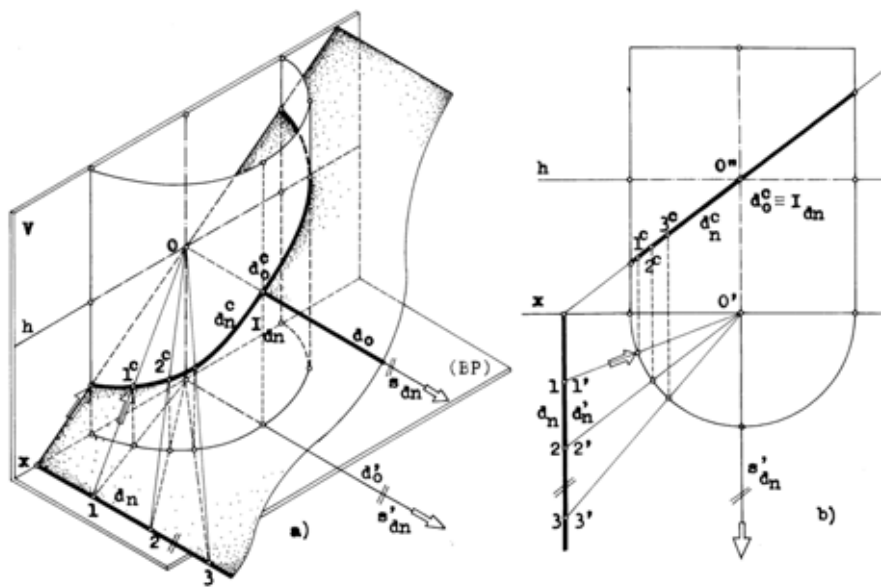


Figure 4: Cylindrical perspective of a line.

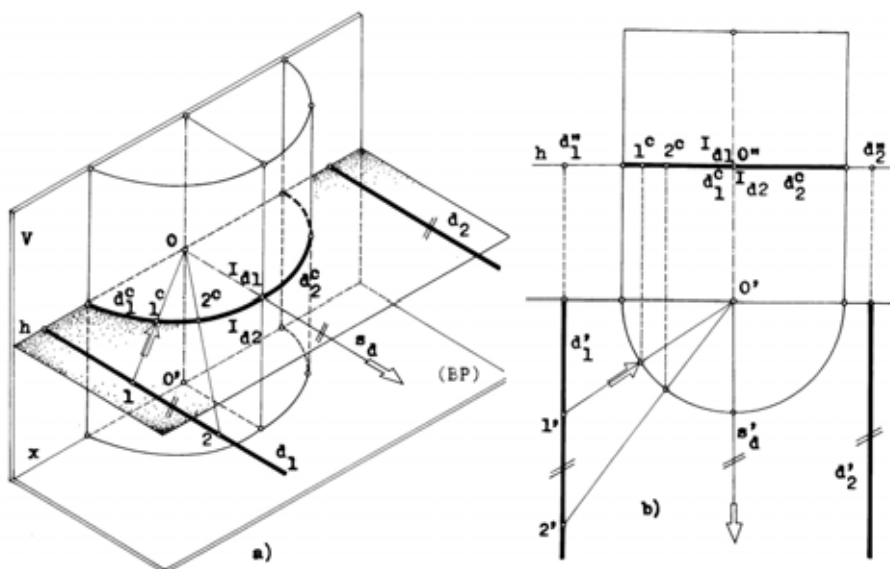


Figure 5: The cylindrical perspective of lines with a specific position.

generatrices of this surface that lie on the basic projection plane V. These infinite points represent the cylindrical perspective Mc and Nc of

points M and N that are infinitely distanced on both opposite sides or directions of line $dh1$. the points of the horizontal dh with the eye fore

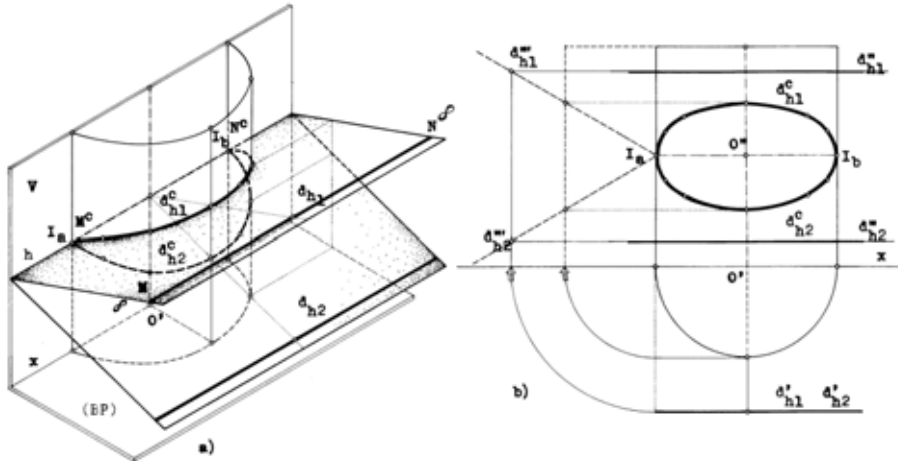


Figure 6: The cylindrical perspective of lines parallel to the basic projection plane (V).

3.1.5.

Figure 7 also presents two lines with specific positions, one of them- line dv , has a vertical position and is parallel with the axis-generatrices of the semicylindrical projection surface. The other line, dh , lies on the basic plane (BP), and passes through the first projection O' of the eye point O and is perpendicular to basic plane V. As we see on figures 7.a) and 7.b) the cylindrical perspective of line dv is determined only for one of its points. Cylindrical perspective lc of point l of this line lies at the intersection of the viewing ray lO with the projection surface. Cylindrical perspective dvc of line dv is determined by the line formed at the intersection of the viewing ray (passing through this line and eye point O) with the semicylindrical projection surface. As it is seen, the viewing ray takes a vertical position there

with the projection surface intersects according to a vertical generatrix. Also the plane point O, has a vertical position. Hence, the perspective dhc of line dh has also a vertical position and lies at the intersection of plane of rays dhO with the projection surface.

3.2. THE CYLINDRICAL PERSPECTIVE AND THE INFINITE POINT OF A LINE WITH ANY POSITION

Figure 8 presents a line d with a spatial position determined through its extreme points A and B. The cylindrical perspective dc of line d is determined by the geometric position of the intersection of viewing rays of its constituent points with the semicylindrical projection surface. Point viewing rays are rays that unite points of a line 1, 2, 3 ... with the eye point O. The characteristics of line d with any

position are: cylindrical perspective of point A and its infinite point. The cylindrical perspective of point A is found at the intersection of viewing ray AO and the projection surface. In fact, the cylindrical perspective Ac of point A is at the intersecting generatrix of the projection surface with the basic plane of projection. Consequently, cylindrical perspective Ac of

point A is on the basic plane of projection V. The infinite point Id of line d lies at the intersection of parallel ray sd with the semicylindrical projection surface. Parallel ray sd is presented by the viewing ray uniting the eye point O with a point of line d that is infinitely distant from the eye basic

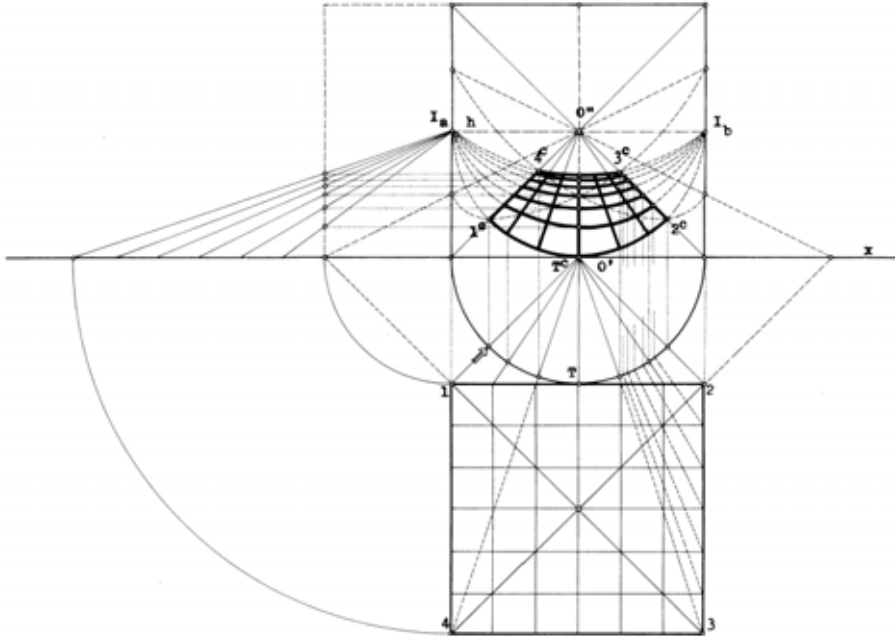


Figure 9: The cylindrical perspective of the quadratic network lying on the basic plane of projection.

plane V. This viewing ray or the parallel ray is parallel to line d and its first projection sd' is also parallel with the first projection d' of line d. The intersecting point of the parallel ray with the projection surface or the infinite point Id of line d lies at the intersection of a vertical drawn from the intersecting point sd' of the base of semicylindrical projection surface with the parallel ray in space sd. This infinite point Id of line d belongs only to one part of a line that lies

before the eye plane V or in the space of real figures. It is evident that all other lines that are parallel with line d share a common infinite point Id.

3.2.1.

Fig.9 present the method of constructing cylindrical perspective of quadratic networks with different positions regarding the basic plane of projection. In figure 9, it is seen that the quadratic network lies on the basic (horizontal)

plane (BP), whereas the corresponding sides of its quadrants occupy a parallel position in regard with the basic projection plane V, respectively with the basic line x. One side of the quadratic network outline touches (tangents) the semicylindrical projection surface on point

T. Therefore, cylindrical perspective Tc of this tangential point lies on the basic line x. Infinite points Ia and Ib for all sides parallel with the basic plane of projection V, lie at the intersection of side parallel rays with the projection

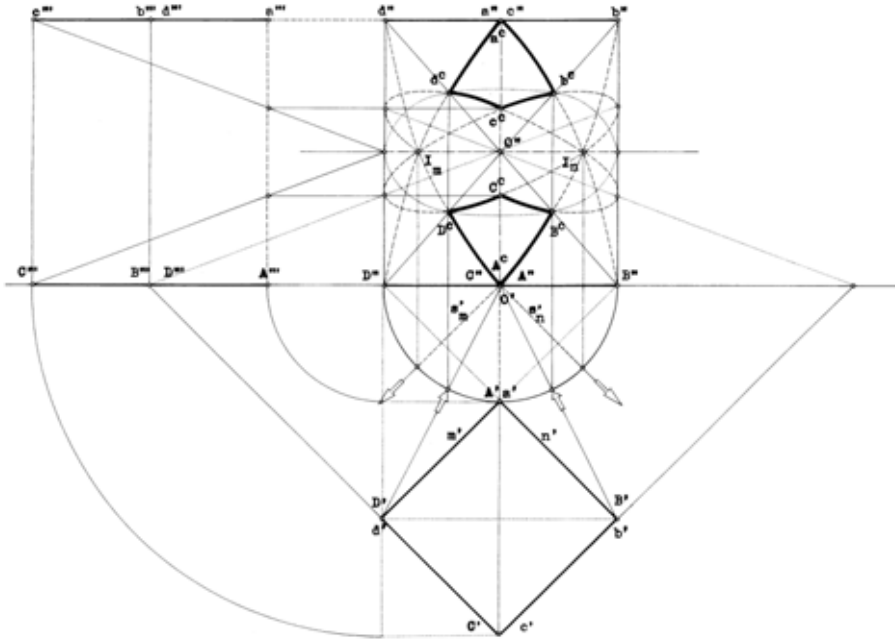


Figure 10: The cylindrical perspective of two horizontal squares.

surface. The parallel ray for these sides is the ray that passes through eye point O and is parallel to these sides; the same coincides with the horizon line h. At the intersection of this parallel ray with the semicylindrical projection surface are found the infinite points Ia and Ib; the same are also found on the horizon line h.

3.2.2.

Figure 10 shows how to determine the cylindrical perspectives of two squares with horizontal position. One of them lies on the basic plane (BP) with one of its vertices touching the semicylindrical projection surface. The other

square lies above the first one in a height Aa. As we know the first projections of these squares are identical. Here, it is also shown how to determine the infinite points Im and In for sides m and n of the square lying on the basic plane (BP) and for the directions- other sides parallel with these sides. Infinite points Im and In are found at the intersection of parallel rays sm and sn with the projection surface. As these parallel rays have a horizontal position and pass through eye point O they also lie on the horizon plane (H), therefore the infinite points Im and In are found on the horizon line h passing through eye point O, respectively

through its second projection O" named as the main point.

3.2.3.

Figure 15 presents how to determine the cylindrical perspective of two quadratic networks, the spatial position of which is shown on a previous figure – fig.13. The squares of the lower quadratic network are shown with interrupted lines whereas those of the upper part are lined. In this fig.15, at the left side of the quadratic network to the basic plane of projection V is continued with the construction of squares which are dotted. While on the right side is continued with the parallel lines – the corresponding sides of the squares, up to the basic line x. It is clear that through these continuations on both left and right sides of the quadratic network it is shown how the cylindrical perspective is constructed.

REFERENCES

- [1] Éric JANTZEN, Traite pratique de perspective, Paris 1983.
- [2] Rudolf SCHMIDT, Lehre der Perspektive und Anwendung, Weisdelberg und Berlin 1983.
- [3] Dr. Fritz REHBOCK, Geometrische Perspektive, Springer-Verlag, Berlin, Heidelberg, New York 1980.
- [4] Dr. Fritz HOHENBERG, Konstruktivna geometrija u tehnici, Beograd 1966.

ABOUT THE AUTHOR

Flamur Doli is an Associated Professor in the fields of Descriptive Geometry, Geometrical Perspective, Traditional housing in Kosovo, Theory and Practice of Restoration. He can be reached by email via address: flamur_doli@hotmail.com; or through the postal address: University of Prishtina, Faculty of Civil Engineering and Architecture, street: Arkitekt Karl Gega, 10000; Prishtina,

Republic of Kosovo. His area of research and expertise is focused on the fields of Geometrical Perspective, Descriptive Geometry, Didactic and its applications. In addition, another field of his field of research and expertise includes studies of traditional architecture. So far, he has published four monograph studies from the field.

GEOMETRIC SHAPES THAT GENERATE ARCHITECTURAL FORMS

Caterina PALESTINI

Dd'A - The University of Pescara, Italia

ABSTRACT: The study, arising from underlying natural geometric shapes that have often influenced planning methods, intends to examine some geometric configurations that derive from sea-shells.

The aim is to interpret the invisible “sinopia” drawings, by the means of instruments of illustration that originate the geometry of these extraordinary natural forms, internally structured by precise patterns. Among these there are some regular helicoid structures; different spirals around which the shellfish build their shells, following a constant compositional rule; a five-pointed star or a pentagon-shaped endoskeleton of echinoderms or sea urchins; and finally, the circular expansion that subtends the geometry of the bivalve shells, symmetrically united by a perfect hinge system.

The fascination with shells, doubtlessly linked to beauty and to the perfection of their shapes, is archaic. This can be proved by the attention that man has always paid them, besides having been nourished by their contents, or having kept them as decoration, distinguishing their decorative features for merely functional or symbolic purposes.

Iconography clearly demonstrates the presence of seashells, or the use of their shape, on different decorative elements or compositions within architectural and artistic fields.

Venus rising from the sea, held up by a shell, was a recurring theme already on Pompeian frescos and on some important paintings, including Botticelli's famous work of art.

Symbolically, the perfection of these naturally created, small structures is enclosed in their shell through which the geometry synthesises the complex mechanism of organic growth. It has inspired many paintings that interpret the purity and the great mystery of the geometric perfection.

This latter is often associated with the maternity of the Virgin, as for example it happens in the apse of the Montefeltro family's Pala di Brera, painted by Piero della Francesca; later it appears many times as motif on ornaments in architecture and decorative art. The quotes about the theme seem to be numerous and variegated, affecting in different ways the story of the art and the architectural planning in which shell geometry can often be found, just as in the case of the famous Guggenheim Museum of Frank Lloyd Wright in New York.

In conclusion, in order to give a brief account of the vast and complicated research that is presented, geometries of various species of gastropods will be analysed and compared by means of different illustrative methods of both traditional and digital design, considering their architectural purposes.

Keywords: Geometry, Nature, Shell, Form, Architecture, Representation.

1. INTRODUCTION

The research to the analysis of hidden geometry into the casing of the shells, examines the characteristics of some specimens, studied to understand the geometric rules and define the design elements that define the formation¹.

The shells, made of calcium carbonate, act as a protective armor of the clam that lives and, since the embryonic Stage, it builds them in layers and in length, through the mantle. In most gastropods the shell, consisting of a single element, act as elicospirale; in some species, defined as primitive, is symmetrical cone shaped much enlarged at the base, with growth direction perpendicular to the axis of winding, as is done in the common patella. In the prevalence of species there is a twisted cone spirally wound around the axis in the form elicospirale or conospirale.

The set of windings determines the coil, consisting of a variable number of turns, among them distinguished from sutures placed to delimit the beginning, suture apicale, and the end, suture abapicale, growth belt. This last one is variously decorated with streaks, grooves and textures that are defined by more or less plastic motifs, is determined by the species to which they belong.

The coil thins in the vertex, consisting of embryonic laps, while on the contrary it widens, at the bottom, in the last perimeter towards the base, which the opening is located generally open to the right, as a result of the direction of rotation, clockwise for the right-handed and counterclockwise for the left-handed, mechanism that can also be reversed or interleave in the initial turns. Within an axial support said columella is the axis of rotation of the coil and can be very compact and adherent appearing without gaps, or be more distanced, resulting in one central hole callednavel, in this case the shell is said perforated. In the end part, at the edges of the opening you may have swellings

and folds, anthropologically known as lips and teeth, which occur during development stages generate evenly spaced bumps called varices; so the growth of invertebrate also systematically defines the ornamentation of the shell that, in its naturalness, follows a precise geometric rule.

2. ANALYSIS OF THE FORMS

Transferring to the images the direct understanding of the complex evolutionary design, which describes the salient features, useful for the subsequent geometric investigation, it can be inferred that the shape of the shell comes from its winding mechanisms, by the number and structure of turns, which in turn depend on the more or less rapid growth of its generating sections.

Regarding this we report the synthesis of certain typological² classifications of recurring forms in the class of gastropods, which are useful for the analysis in question; derived from studies on the subject³.

It showed, according to the type of winding, evolved forms with winding conospirale, among them tangential and clamped, facing outward; involuted forms wound on themselves with the same mechanism facing the interior, equipped with a visible navel; convoluted with the last turn it clings to, summing up on itself with the previous turns. The identification continues by similarity with the main forms such as cones, including the conical shell that reminds the shape of the cone, flattening on the upper part; the biconical shell with moderately high spire that duplicates the shape; the citroconica always conical with convex sides and coeloconic with concave sides.

Further similarities are identified with objects, from which the nomenclature of fusiform shells with long and a tapered base; the turricolate high spiers that resemble towers with narrow and elongated spira; the troncheiform with acute spira and flat base; the

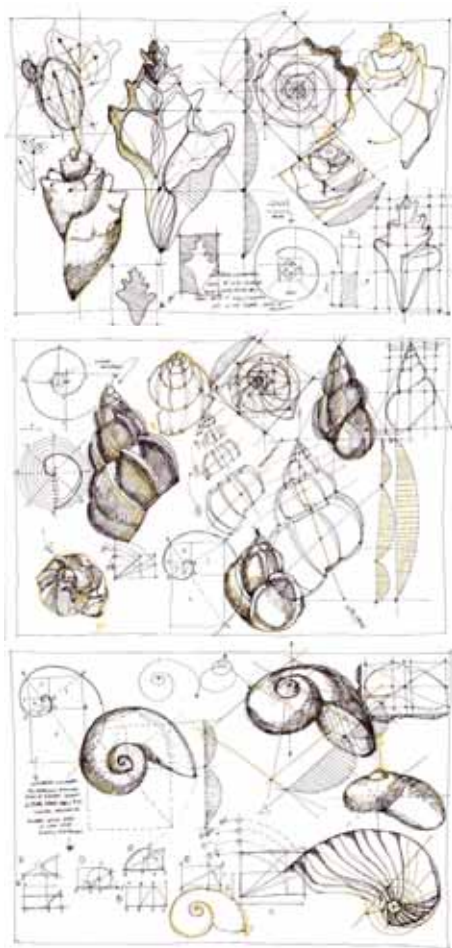


Figure 1: Graphical analysis of the forms.

turbinated coils in the shape of a spinning top with a large spira, rounded turn and convex base; the *discoid* disc-shaped, convoluted or involuted form; the *globular* or *barrel round shaped*. We add then the similarities with other natural forms, among them, just to mention a few, the olive, the pear, the egg, the ear. It is clear that the complexity of the topic, which is the subject of numerous multidisciplinary studies, would require a broad

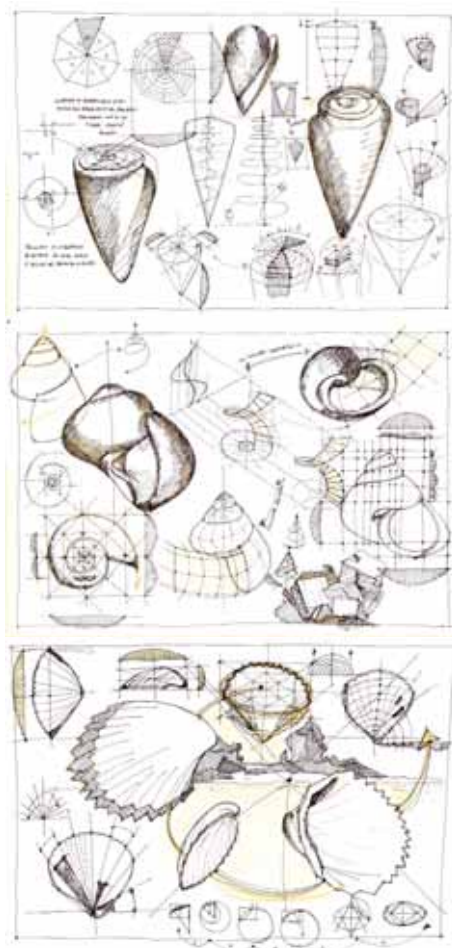


Figure 2: Graphical analysis of the forms.

der study compared to this exhibition, dedicated to issues related to the representation and geometric analysis of these elements. The aim of the research was to synthesize the identification characters of the type considered, of which the proposed digital studio models, reconstructed according to the analysis of geometric generating. The analysis took into consideration some species belonging to the category of *Bival*



Figure 3: Classification and analysis 3d.

ves, formed by two shells joined together by means of a hinge, determined by an ingenious system of such protrusions, *teeth*, cavities, and *dimples*, which fit together allowing the opening and closing movements, facilitated by an elastic ligament nature of the cornea. Unlike *Gastropoda* that need a movable Shell to make easy movements in search of food, the *Bivalves* are sedentary filter feeders and use their protective armour to defend

themselves from predators, submerging with speed in the sand. The wedge shape promotes infiltration in soft bottoms, the structure set up for this purpose, can have identical shape valve defining it *equivalve* or slightly different, *inequivalve*; in each of them we finds a starting point, *umbo*, on which lays the dorsal of the shell and the ventral on the opposite side.

During growth, the shells alternate development stages and moments of stasis, this succession generates on the outer surface concentric striations, the so-called *growth striations*. The decoration is similar to that of the *Gastropoda*, may have in addition other streaks and coastline; placed in respect to the origin, in a concentric or radial manner, respectively defined *concentric sculpture* and, or *radial*. In some cases, in fact, both exist, along with any other eventual spines, tubercles and lamellae, that the species has raised for functional reasons and camouflage, with respect to the seabed or purely ornamental. We cannot miss a hint to the family of cephalopods, as it belongs to the extraordinary *Nautilus* the only species to have a true external shell which employs it as a hydrostatic stabilizer. In reality these predatory animals, basing their survival on movement speed, have left the shelter, preserving rudimentary parts of domestic support, as in the case of the common cuttlefish. Among these the *Spirula* containing the thin shell, flat spiral, in its body, sharing with the *Nautilus* the feature of gas filled chambers, used for the flotation. Another case, in the same category is that of the *Argonauta* that among his tentacles when needed generates, using limestone secretions produced by females, a thin shell that spawns and abandons it at when they disclose. The fragile-looking parchment shell is streaked with coasts that radiate from the central core, based on a geometric progression comparable to a logarithmic spiral that notoriously distinguishes that famously characterizes harmonic completeness of the *Nautilus*. The shell, symmetrical and bilateral

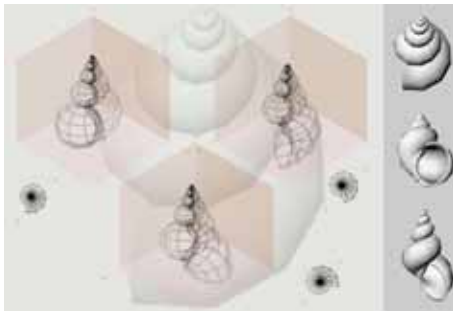
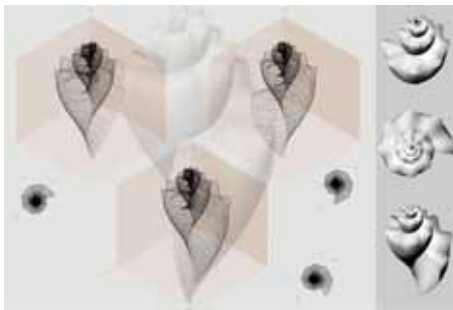
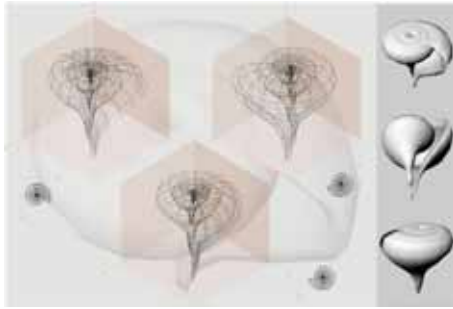


Figure 4: 3d models.

with a grand opening but without a navel, also in the perfect *Nautilus*, it appears light, but structured in the compartmental chambers, connected by a thin tube; these areas distinct from one another by septa cross, enable the mollusk that lives curled up in the last environment, to check the dive and float system.

A complete balance between form and function adjusts the hydraulic engineering and its

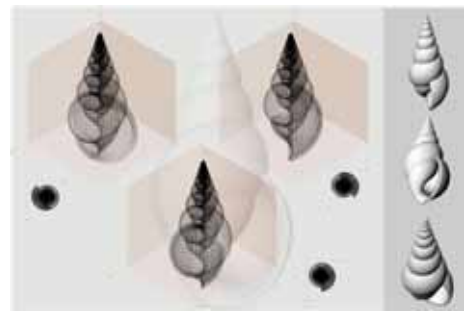
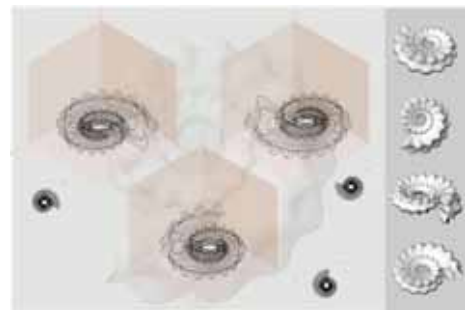
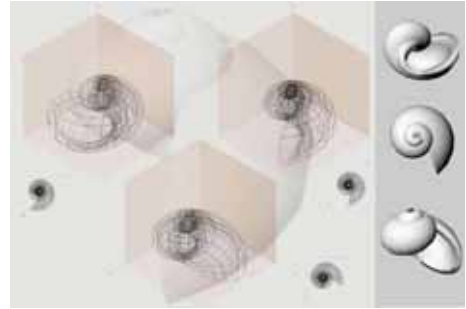


Figure 5: 3d models.

geometry, underlined also in the ornamental design of the external surface which, on a white shell, it outlines reddish streaks, which mention rhythmical succession of the interior. In the two examples of cephalopods, similar by typology to the *discoid's*, the generating figure rotates generating through a perpendicular plane to the axis, increasing its size in geometric progression; unlike in most gastropods, such as the *turbinates*, it outlines a

helical path in respect to the development axis, which determines a curve defined *elicospirale*⁴.

It should be noted that the investigation on the shells, while experiencing different spirals forms, evaluated in the geometric reconstruction of the models, remains unchanged the basic principle of their growth, which is accomplished by expanding on their own in a single direction, without changing the initial configuration, according to a proportion that stays unchanged⁵. This mechanism related to the concept of *gnomon* which by definition has the similarity between the final figure and the original one, despite others are interposed, it has been encountered based on analogy and extended to some forms of natural growth and among these shells, where each extension adds a part gnomonic to pre-existing and the overall structure⁶.

3. RECURRENT GEOMETRIES

Based on these preliminary assessments, the survey conducted on a selected number of examples, organized by families and membership features, has tracked down the generating curve design inscribed in the shells, and then develop three-dimensional models for study purpose. The recurring forms are the results of verifications carried out, as it was intuitable, have appeared indicative the Archimedean spiral and logarithmic. The first is obtained by drawing from a point of origin a circumference continuously developed, increasing the radius in a consequential mode; the logarithmic spiral or equiangular follows a geometric progression, proceeding in a infinite way to the external and the internal of its shape. The properties of this curve defined in the seventeenth century, first by Cartesio and later by Bernoulli that was fascinated, actually has previous derivations linked to Fibonacci numbers, in which, as is well known, each digit is the sum of the previous two, as clearly evident from the geometric construction of squares, that are arranged in systematic progressive succes



Figure 6: Classification and analysis 3d.

sion generating the logarithmica⁷spiral. The two concerned spirals, defines both open curves, one is a uniform type and more static which is identified and drawn in points, starting from an extreme point of a radius than moves smoothly around its extreme; the other, more dynamic is built traditionally, with a compass connecting the successive arcs of circumference from a square core that expands into a progres

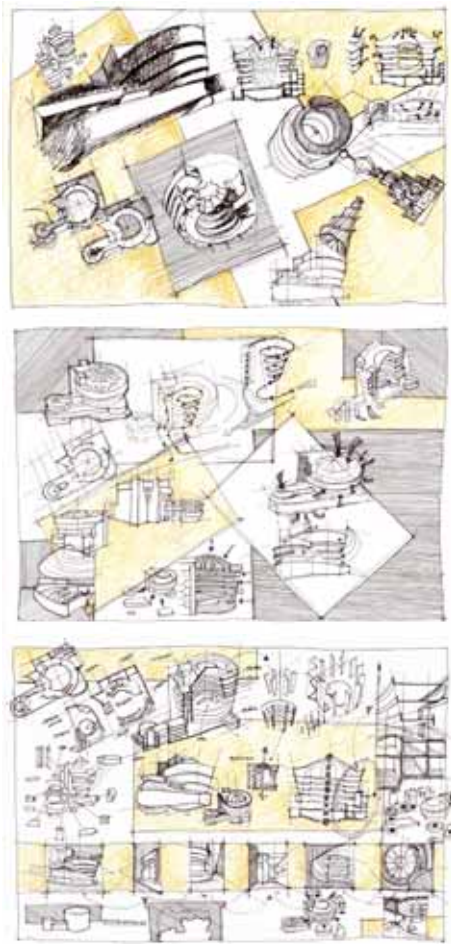


Figure 7: Analysis Guggenheim Museum in a form similar to a shell.

sion of golden rectangles, which generates the golden spiral⁸.

The shells as described for growth and traction evolve by describing a conical helical that moves in three major forces: a circular motion; an expansion outwards, with *evolved* forms, or internal, *involute*; and traction in the direction of the vertical axis. When the point has approached the circular motion, is

corresponding to the origin, but at a higher proportion and more enhanced, it is interesting to note that the enhancement is called step and a full turn is defined as spira, with a nomenclature corresponding to the one adopted for the gastropodi classifications.

The illustrations⁹ accompanying the text, related to the analysis of geometric matrix of shells, take into consideration the topic from real forms, designed initially with traditional graphic surveys, to arrive at the realization of three-dimensional models, created with the help of digital representation¹⁰.

The illustrations, highlighting the paths and generating geometrics showing the following method, specifying through a graphic analysis, the interest towards inspiring shapes for architecture. It could not, therefore, go unnoticed a shell that as the name reveals its connection with the edificatory art, *Architectonica perspectiva*, of the *Architectonicidi* family, characterized by a *swirled shape* with uniform laps that remind of the ramps of a spiral staircase, wrapped around the deep navel, *seen* from a bottom view they determine a singular perspective effect. Still on the subject of analogies linked to build the *Thatcheria mirabilis*, of the *Turridi* family, is characterized by a winding *flat spiral* which generates a web like membrane and rising above determine a cusp that reminds us of the Borromini lantern of s. Ivo alla Sapienza; continuing among many analyzable, was chosen the structure of *Epitonium scalare* is a small diaphanous shell, outlined by rounded coils without suture loops, connected by thin coastline and detected that mark crosswise the twisted annular perimeters. Then there are the *Turricolate* with different species of *Turritelle* distinct by a slender appearance that makes them comparable to soaring towers, comparable in form to the pinnacles. The *Conidia* are also being reviewed, consisting of numerous examples that in various forms already mentioned, and with a completely flat or elevated spira, present even on the smooth surface of the shell geometric de-

signs that are repeated systematically with modules and constant angles, as well as their compositional structure.

In addition to direct analogies between the shapes of shells and architectural ones, also expressed by names, there are many relationships with design inspirations¹¹ arising from, in particular connected with organic architecture. It is an emblematic example of the famous Frank Lloyd Wright's work, specifically revisited through the design, which configures the Guggenheim Museum in a form similar to a shell, by internally going through a spiral path.

4. CONCLUSIONS

To delimit the result of the extensive search that might continue with several other notable examples, the contribution given here, wants to set up just a small block that stands more as a beginning than an end, aware of the vastness of the topic and the much broader research conducted on the subject¹²; the first concern from which flow the above considerations and graphic analysis reported in the text and in the summary synoptic table.

REFERENCES

- [1] Wolfgang Goethe considered the spiral as a mathematical symbol of life and of evolution, the theory that found objective evidence in polynucleotica DNA that develops duplicating with a cyclical spiral movement, also present in many species of gastropods.
- [2] Cfr. B. Sabelli, *Conchiglie*, Milano 1980.
- [3] For news related to the characteristics of mollusks and their classifications were consulted several atlases and specific texts, among them see. C. and F. Ghisotti Ashlars, *Shells*, atlases, Florence 1998 Nature; S. P. Dance, *Shells*, and. italiana a cura di a. Daga, Milan 2003; *Phylum mollusca* a cura di g. Simone and Jessica r., laboratory of Paleontology, www.etna2000.com
- [4] Cfr. D'Arcy W. Thompson, *Growth and formation*, Torino 1969, pp. 191-222-

- [5] Cfr. M. Holt, *Mathematics in art*, London 1971.
- [6] This mechanism has been masterfully described in Thompson's volume that constitutes an indispensable point of reference for research on the geometry of natural forms. See D'arcy w. Thompson, *growing cit.*, p. 201-209.
- [7] Cfr. M. Docchi, *Disegno e analisi grafica*, Bari 1985, vol. I, pp. 41-42.
- [8] Cfr. A. Montù, *Natura e Geometria*, Milano 1970.
- [9] The research on representation of shells, real shapes to digital models, launched in 2003 partnered Gianna Cieri, pupil student of a course degree in design, which drew up the graphics on the topic of study, coordinated by me
- [10] Cfr. R. Migliari, *Disegno come Modello*, ed. Kappa, Roma 2004.
- [11] Cfr. I. Calvino, *Lezioni Americane*, ed. Mondadori, Milano 2000
- [12] Cf. C. Palestini, *geometries: sinopie of the form*, in *the geometry between teaching and research*, Atti del Convegno Internazionale, Università degli Studi di Firenze, and. Area, Firenze 2008, pp. 379-384.

ABOUT THE AUTHORS

Caterina Palestini, architect, associate professor of design at the Department of di Architecture of the University "G. d'Annunzio" Chieti-Pescara.

THE GEOMETRICAL APPROACH TO CALCULATION OF BOUNDARY LAYER BETWEEN FLOWS

Olena GUMEN¹ and Viktor MILEIKOVSKYI²

¹National Technical University of Ukraine “Kyiv Polytechnic Institute”, Ukraine

²Kyiv National University of Construction and Architecture, Ukraine

ABSTRACT: An interaction of concomitant or opposite flows is used in a technique. In particular, the initial part of current is formed by interaction of flow running from an opening or a slot and the ambience. This paper presents an approach to calculate such flows using a theory of professor A. Tkachuk based on method of features. We accept large-scale vortexes as foreign bodies (features). An idealized scheme of a turbulent boundary layer uses the sheet of large vortexes that symbolically enlarged until the touch.

For the flows running in the same directions we obtain good coincidence with Abramovich's theory. For the flows running in opposite directions we obtain solution different from Abramovich's theory. At equal velocities the new solution is symmetrical that is more acceptable than the Abramovich's theoretical asymmetrical solution. We show problems of experimental prove. But we perform CFD experiments with flows similar to our. And we obtain good coincidence with our calculations.

Keywords: Turbulent boundary layer, turbulent mixing, CFD.

1. INTRODUCTION

An interaction of concomitant or opposite flows is used in a technique. In particular, an initial part of a current is formed by interaction of flow running from an opening or a slot and the ambience.

The laminar boundary layer can be precisely described by Navier-Stokes equations. But for the turbulent one equations are more difficult and have less precision. The development of turbulent boundary layer is accepted as a result of action of turbulent pulsations in flows [1,3,4,5,6,7,9]. In such context the boundary layer theoretical descriptions may use experimental coefficients. As a results we need to perform difficult and expensive experimental researches.

Today the most popular method of turbulent flow calculation is CFD simulation in 2D or 3D virtual model [1,4,5,6]. It is simpler and cheaper than real experiments but the methodology is the same. We need to plan virtual experiments, build 2D/3D virtual model (we save time and money for materials and work), perform experi-

ments (some time saving because a numerical experiment is quicker), approximate the results. Also we need to validate the results by performing some real experiments. Most of modern CFD models [1,4,5,6] contain a lot of experimental coefficients and we can't prove that the coefficients are task independent.

The mathematical (or geometrical) description of flows did not loose topicality [3, 9]. It can give possibility of direct optimization, obtaining engineering dependencies without additional experimental planning and approximation. We can easily calculate parameters at any input data without additional simulation.

Professor of Heat Gas Supply and Ventilation Department of the Kyiv National University of Construction and Architecture A. Tkachuk [10] proposed a theory of turbulent boundary layers by the method of features. A flow is examined as a flow of ideal liquid. Small vortices plays the role of "features" as foreign bodies. Tangential velocity component discontinuity in small-scale turbulent structure modelled as a vortex sheet. The vortices periodically escape the sheet by

Magnus forces. But vortex generation patches the sheet immediately. Thus the action of vortices maybe directly described instead of additional concepts usage.

The turbulent boundary layer between flows (fig. 1) has the same vortex sheet with two differences: a) the vortices are large-scale, b) the boundary layer enlarges. These facts causes a possibility of geometrical analysis.

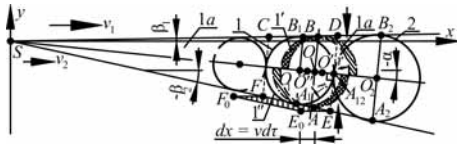


Figure 1: The boundary layer

2. GEOMETRICAL MODELLING PROBLEMS

Before we start analysis we need to analyse some problems and concepts.

2.1 Geometrical compatibility problem

The vortices roll on flows 1 and 2. Both flows are flat with even velocity profiles. Corresponding velocity projections to a direction of flow 1 are v_1 and v_2 . Let us consider $v_1 > v_2$ and $v_1 > 0$. The vortices move with constant velocity with the x -component

$$v = (v_1 + v_2) / 2. \quad (1)$$

If vortices do not enlarge during development there are no collisions during movement. But if the boundary layer enlarges next vortices impose upon previous. This causes vortices deformation (to ellipse) stopped by a) disintegration to smaller vortices or b) integration to next larger vortices. After that previous vortices will catch up with next ones repairing the macrostructure (fig. 1). It cause some layer randomization.

2.2 Averaging problem

Most of technical tasks require averaging of turbulent flow parameters. We can perform it by two ways: a) calculation of all momentary macrostructure states and perform averaging or b) drawing the most common (average) state of

macrostructure and analyse it. Second way is more simple so we accept it. The average state between deformed vortices and round vortices at the distance is near to shown on fig. 1. The vortices are at the contact. The contact may be modelled by very small intensive vortex caused intensive flow mixing.

3. GEOMETRICAL ANALYSIS

3.1 Velocity profile

The geometrical analysis of velocity profile was performed in work [8]. We accepted that a vortex rolls by flows as a solid wheel. In a part between vortices ($B_1A_{12}B_2$ at fig. 1) there is only an injection of flow liquid (gas) into the boundary layer in the direction of an axis y normal to an axis x . The axis x aligned in the direction of greater flow velocity v_1 . The x -component of velocity in this part of boundary layer is equal to corresponding flow velocity because of inertia. The same concept we may use to calculate boundary layer enlarge.

3.2 The boundary layer enlarge

For $v_2 > 0$ the precise equations are published at proceedings of a conference "Environmental engineering" in Lithuania [2]. In this work we will try to simplify them without a precision loose. We will evaluate an error of simplifications.

Let us consider [2] two round touched vortices 1 and 2 (fig. 1). In the vortex 1 we build a diameter A_1B_1 normal to line O_1O_2 . The point A_1 is at the boundary with the flow 2, the point B_1 is at the boundary with the flow 1. In the vortex 2 we build a similar diameter A_2B_2 . After an infinitely small time interval dt the vortex 1 occupies a position 1a with radius r and a centre O . The points A_1 and B_1 translate to A and B .

We accept the linear boundaries of the layer. Angle between axis x and O_1O_2 we call $\alpha = \text{atan}(\Theta_0)$ with positive direction into the flow 1. In most of cases $\alpha < 0$. $\beta_i = \text{atan}(\Theta_i)$, $i = 1, 2$ are angles between the boundary with the flow i and the axis x . The positive angle direction is into the flow 1, so $\beta_2 < 0$. All angles are in the range from minus $\pi / 2$ to $\pi / 2$.

Thus all vortices are homothetic so [2]

$$r = \theta x, \quad (2)$$

with constant coefficient θ of vortex radius growing.

We circumscribe the vortex 1 by a trapezoid $CDEF$. The points C, D are at the boundary with the flow 1 and E, F are at the boundary with the flow 2. The lines CF and DE touch the vortex 1. It has medial line and high equal to diameter of the vortex 1. So the area is $S = 4 r_1^2 \approx 4 r^2$.

Let us consider contour (not shown on fig. 1 to avoid overloading) with the radius r concentric to the vortex 1. We move both half of this contour along the line O_1O_2 with the x projection of velocity equal to the corresponding flow velocity during infinitely small time interval dt . So, we obtain arcs 1a and 1b. The same arcs may be obtained by moving half of the contour 1a along O_1O_2 by the distance along the axis x equal to $(v_1 - v) dt$ in the direction according to the difference sign. Because of this areas of net-hatched figures using the equation (1)

$$S_1 = S_2 = r (v_1 - v_2) dt / (2\cos(\alpha)). \quad (3)$$

The macrostructure of the boundary layer shown on fig. 1 can be only if both net-hatched areas S_1 and S_2 are consumed by the layer. In different case the vortex will push out or push some particles between vortices. If $v_2 < 0$ the net-hatched and equivalent line-hatched area belong to the flow 2 must be exchanged. In this case we can translate the line EF with the velocity v_2 at the distance $v_2 dt$ during the infinitely small time interval dt . We obtain new line E_oF_o . The area S_3 of parallelogram FEE_oF_o will be excluded from the boundary layer during the time dt . If $v_2 > 0$ then $S_3 = 0$.

As a result we can perform geometrical analysis of flow balance. We can present the flow balance of boundary layer as an area balance: during the infinitely small time interval dt the growth of parallelogram $CDEF$ area dS_{CDEF} is equal to sum of infinitely small areas S_1, S_2 and S_3 :

$$dS = k (S_1 + S_2) + S_3, \quad (4)$$

where factor $k \leq 1$ takes into account a flow

mixing near to the point A_{12} of the vortices touching. This factor will be found analytically.

So, the task of macrostructure analysis simplifies to the task of characteristic areas calculation. The equation (4) appears to be differential. But all areas in the right side are proportional to dt . The left side is proportional to dx . Using $dx/dt = v$ the equations (1...4) will stay algebraic. At $v_2 > 0$ for small angles α we can put $S_3 = 0$ and $\cos(\alpha) \approx 1$. Solving the equations (1...4) gives a simple formulae, simpler then [2]:

$$\theta = (k/4) (1 - (v_2/v_1)) / (1 + (v_2/v_1)). \quad (5)$$

At $v_2 = 0$ the value of θ is the greatest. The angle $|\alpha| < \text{atan}(\theta)$. In different case all boundary layer will be immersed in one flow. So by the equation (5) $|\alpha| < \text{atan}(1/4) = 0.24$ and $\cos(\alpha) < \cos(0.24) = 0.971$. So the error of eliminating $\cos(\alpha)$ is less than 3%. It is acceptable because the experimental error is approximately 5%. For $k = 1/2$ $|\alpha| < \text{atan}(1/8) = 0.124$ and $\cos(\alpha) < \cos(0.124) = 0.992$. The error is less than 1%. The factor θ is independent on an angles of the boundary layer. So we can avoid an equation system solving process at this stage. At $v_2 < 0$ the task is more difficult. We need to take into account S_3 and can't eliminate $\cos(\alpha)$:

$$\theta = \frac{\frac{k}{4} \left(1 - \frac{v_2}{v_1}\right) - \frac{\Theta_2 (v_2/v_1)}{1 + \Theta_2 \Theta_0}}{1 + (v_2/v_1)} \sqrt{1 + \Theta_0^2}. \quad (6)$$

Solving triangles and using the equation (2) we can write the angles relations as:

for $v_2 < 0$ (and also for $v_2 \geq 0$):

$$\Theta_i = \frac{\Theta_0 \sqrt{1 + \Theta_0^2} \pm \theta}{\sqrt{1 + \Theta_0^2} \mp \Theta_0 \theta}. \quad (7)$$

for $v_2 \geq 0$ we can use very simple equation:

$$\Theta_i \approx \Theta_0 \pm \theta; \quad (8)$$

The top signs are for $i = 1$ and bottom – for $i = 2$. The equation (7) is correct for $v_2 \geq 0$. But in this case the greatest value of $\Theta_0 \theta$ using the

equation (5) will be at $v_2 = 0$:

$$\max(\Theta_0 \theta) < 0.24 \cdot 0.25 = 0.06.$$

But the square root equal to $\cos(\alpha)$ is near to one. So the 6% error of eliminating the denominator in equation (7) is comparable with the experimental error at $k = 1$. At $k = 1/2$ the error will be less than 2% because

$$\max(\Theta_0 \theta) < 0.124 \cdot 0.125 = 0.016.$$

So the equation (8) is usable. It is very simple so it can be substituted in any equation without any effort. The value of θ is known from equation (5).

To find Θ_0 we need additional equation(s) dependent on flow types. If both flows are free (free boundary layer) we can use a momentum equation in projection to axis y . The equation (12) in work [2] is acceptable for any sign of v_2 . It can be transformed to the following form using equations (7) – (11) in [2] that is independent on v_2 sign:

$$\left(5\Theta_1 - \left(4 + \frac{v_2}{v_1}\right)\theta\right)\left(4 + \frac{v_2}{v_1}\right) + \frac{\rho_2}{\rho_1}\left(5\frac{v_2}{v_1}\Theta_2 + \left(4\frac{v_2}{v_1} + 1\right)\theta\right)\left(4\frac{v_2}{v_1} + 1\right) = 0. \quad (9)$$

For $v_2 \geq 0$ we substitute the equation (8) to (9) and obtain a solution:

$$\Theta_0 = \theta \frac{\left(\frac{v_2}{v_1} - 1\right)\left(\frac{v_2}{v_1}\left(4\frac{\rho_2}{\rho_1} + 1\right) + \frac{\rho_2}{\rho_1} + 4\right)}{5\left(\frac{v_2}{v_1}\left(\frac{\rho_2}{\rho_1}\left(4\frac{v_2}{v_1} + 1\right) + 1\right) + 4\right)}. \quad (10)$$

We avoid solving of equation system for $v_2 \geq 0$. First we use the equation (5) to calculate θ . After that we find Θ_0 by the equation (10). And finally we can find Θ_1 and Θ_2 using the equation (8). Only for $v_2 < 0$ we need to solve the equation system (6), (7) – two equations for $i = 1, 2$ – and (9).

To find value of k we need to compare a semi-restricted boundary layer (one of the flows is restricted along it without an injection possibility) and the free boundary layer [2]. An example of semi-restricted boundary layer is the layer between a current initial part and ambience. If the flat current starts with even velocity profile it contain a core with constant velocity and linear boundaries that interacts with ambience. At the start the boundary layer has to small sizes to take into account restrictions. But during development both boundary layers prevent each other to eject a liquid or gas from the core. So at one side of each boundary layer ejection will decay to zero that is equivalent to free boundary layer at $k = 1/2$. Because we could not observe some changes in boundary layer development we can use $k = 1/2$ for free and semi-restricted layers. For semi-restricted boundary layer we need to replace the equation (8) or (9) by another one to describe the nature of restrictions. It is impossible to cover all possible restrictions by common equation(s).

4. THE RESULTS VALIDATION FOR FREE BOUNDARY LAYER

The results are built in Figure 2. At $v_2 \geq 0$ we obtain good coincidence with Abramovich's theory [7]. This theory is more complex and did not provide simple engineering equations such as (5), (8) or (10). At $v_2 < 0$ there are too few experimental results [7] only up to $v_2 / v_1 > -0.3$. But the Figure 2 shows that the experimental data is closer to the calculation results by the equations (6), (7) and (9).

At $v_2 < -0.3$ the absence of experimental data caused by the problems with interaction between a source of the flow 2 and throttled boundary layer. It causes very difficult resulting flow running away from the mixing boundary layer. So we obtain no velocity at the side of flow 2 near to flow 1. But we can workaround the problem. As we show above the starting part of the semi-restricted boundary layer is similar to free. Also we can use the starting part of a restricted layer (both flows are restricted).

At Figure 2 we show results for CFD simulation of two opposite flows in a short square pipe (Figure 3). The pipe length is equal to high and width of it's section.

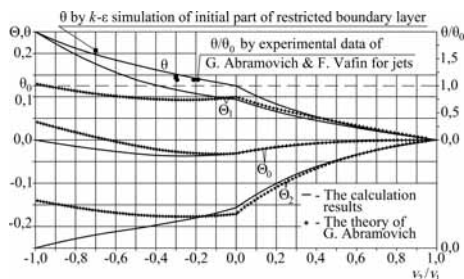


Figure 2: The results of calculations

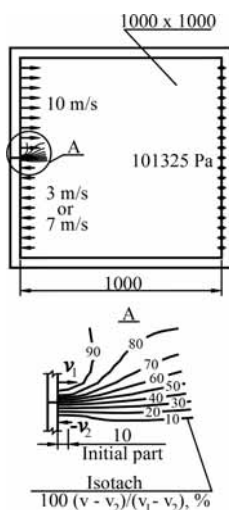


Figure 3: Scheme of duct CFD model

The initial part is linear and has a length near to 1% of the pipe length. The results are similar to obtained by the equations above. So we can indirectly prove these results. Also the results give symmetry of boundary layer at $v_2 = -v_1$, that is more considerable.

5. CONCLUSIONS

1. We propose a scheme of turbulent boundary layer macrostructure between flows. The scheme shows an averaged state of the boundary layer during its development.
2. Geometrical analysis of the boundary layer using a momentum equation (only for free boundary layer) gives simple easy engineering equations to calculate boundary layer between

flows with the same direction and gives a system of four equations for opposite direction of the flows.

3. For the same directions of flows the results are similar to Abramovich's theory. For opposite direction obtained results are closer to experimental data. For high velocity ratio we use the initial part of restricted boundary layer to prove the results. The results are close to CFD simulation data. So we prove the new approach to calculate turbulent boundary layer between flows.

6. PROSPECTS OF THE RESEARCHES

The next task is to solve a current developed in a flow. The analytical description was finished but we obtain the similar problems with opposite flow. The approach gives us a critical velocity ratio of the current destroy. The CFD simulation proves the analytical results and shows the destroy process. Now we perform visual experiments.

ACKNOWLEDGMENTS

The authors thanks to our famous scientist and teacher, professor of Kyiv National University of Construction and Architecture Andrii Tkachuk (1928-2002), founder of this research direction in the University.

REFERENCES

- [1] E. Elhadi, L. Xiaosong and Wu Keqi. Numerical simulation of 2D Separated flows using two different turbulence models. *Pakistan Journal of Applied Sciences*, 2(12): 1057-1062, 2002.
- [2] V. Mileikovskyi. The Macrostructure Analysis of the Turbulent Mixing Boundary Layer Between Flows with the Same or Opposite Direction In *The 9th International conference "Environmental Engineering" (Vilnius, Lithuania, 22-23 May 2014). Selected papers.* eISSN 2029-7092 / eISBN 978-609-457-640-9. Available online at <http://enviro.vgtu.lt>, 2014. Direct link: http://leidykla.vgtu.lt/conferences/ENVIRO_2014/Abstracts/6/273.html
- [3] J. D. Ocon, L. Jirkovsky, R. L. Leon and A. C. Muriel. Theory derived law of the wall for parallel flat-plated turbulent flow. *CFD*

- Letters*, 4(3): Pages 93-101. – September 2012
- [4] M. Sjeric, D. Kozarac and R. Tomic. Development of a two zone turbulence model and its application to the cycle-simulation. *Thermal Science*, 18(1), 1-16, 2014.
- [5] I. Stankovic and B. Merci. LES-CMC simulations of a turbulent lifted hydrogen flame in vitiated co-flow. *Thermal Science*, 17(3): 763-772, 2013.
- [6] S. Tamini. Simulation of turbulent flow using FEM. *International Journal of engineering and technology*, 2(8): 1358-1363 , August 2012.
- [7] Абрамович, Г.Н. 1960. Теория турбулентных струй [The turbulent currents theory]. Москва: Государственное издательство физико-математической литературы [State physics and mathematics literature press]. 715 S.
- [8] О.М. Гумен, В.О. Мілейковський. Геометричне обґрунтування профілю швидкості примежового шару між зустрічними потоками In Гідравліка і гідротехніка [Hydraulic and Hydrotechnic]: Науково-технічний збірник, pages 11-18. Вип. 66.– К.: НТУ, 2012.– 120 с.
- [9] Г.П. Скребков, Фёдоров Н.А. Интегральная и локальная величины коэффициентов турбулентного профиля скорости [Intergal and local values of coefficients of turbulent velocity profile]. *Вестник МГСУ [Bulletin of MSCU]*, No4: pages 201-208, 2013
- [10] А.Я. Ткачук, Е.С.Зайченко, В.А. Потапов, А.П.Цепелев. Системы отопления. Проектирование и эксплуатация [Heating systems. Design and operation]. Киев: Будівельник, 1985. 136 Pages.
- E-Mail: gumens@ukr.net .
2. Viktor Mileikovskiy. PhD., docent of Heat, Gas Supply and Ventilation Department of Kyiv National University of Construction and Architecture.
E-Mail: mileikovskiy@gmail.com .

ABOUT THE AUTHORS

1. Olena Gumen, Doctor of Science, professor of Department of Descriptive Geometry, Engineering and Computer Graphics of National Technical University of Ukraine “Kyiv Polytechnic Institute”

GEOMETRICAL ASPECTS OF CITY SKYLINE – TALL BUILDING ANALYSIS

Klara CZYŃSKA

West Pomeranian University of Technology Szczecin, Poland

ABSTRACT: Virtual 3D city models and digital analytical techniques may significantly facilitate the process of studying a city structure. They enable imaging and conducting of advanced analysis of city skyline. They can be applied to tall buildings phenomenon which became one of the most important challenges for contemporary European cities. On the one hand, we attempt to maintain the unique historical values of cities and on the other, we tend to create a contemporary image of a city with tall buildings as a symbols of modernity. Therefore, in order to protect valuable urban and architectural development, location of tall buildings should be determined on the basis of prior in-depth analyses. The article presents geometrical techniques based on digital processing virtual city models: methods of visual impact range, view range and view angle analyses. Those methods have been successfully used in planning process as a basis for defining guidelines for erecting new tall buildings and protecting old city skylines.

Keywords: 3D virtual city models, 3D analysis, tall buildings, city complexity, urban planning.

1. TALL BUILDING IN EUROPE AND CULTURAL HERITAGE PROTECTION

Landscapes of European cities result from centuries of evolution taking place in urban systems. A novelty, from the point of view of historical pace of development, are new tall buildings. In the 21st c., tall buildings have become a fixed element in landscapes of many European cities. This development is increasingly popular in various countries, including Scandinavia which up until recently preferred low rising buildings, for example Malmo and Stockholm. Following footsteps of American and Asian cities, the largest metropolises of the old continent transform their historical landscapes. A skyscraper is a generally accepted attribute of modernity and grand cities, and interpreted by many people as an attribute of modernity, prestige and economic advancement. A cluster of tall buildings is a powerful symbol of a modern city, a fascinating view attracting tourists in large numbers [25].

However, developing of tall buildings in-

volves major threats to landscape cohesion and integrity in European cities. Specific architectural and urban arrangement reflected in the silhouette of a city is an important part of the protected cultural heritage. Due to their broad visual impact range, tall buildings frequently induce unfavourable and unplanned interaction with historical buildings. They diminish the influence of primary architectural dominants (e.g. towers of churches and town halls) as regards their role in the overall composition. For this reason, plans to develop tall buildings trigger conflicts and controversies. In Europe, more than on other continents, it is believed that a tall building should have particular architectural and spatial features [18]. A tall building is considered to be an icon in terms of its nature and position in the urban structure. In order to proceed with an objective discussion on the role of tall buildings as such, we need to develop a methodology for assessing and planning such buildings in a city landscape. We need to fully document the future, planned vis-

ual impact to formulate reliable and competent planning guidelines and strategies for landscape protection and development.

In many European cities, we can find effects of wrong decisions concerning the location of tall buildings. One of examples is the panorama of London as seen from the river of Thames. The panorama shows a lively expansion of various architectural forms. Several tall buildings are situated not only in the City or Canary Wharf, but they are also scattered in various parts of the city. Due to the above, in some important views, elements of historical cultural heritage, such as St. Paul's Cathedral, lose their significance (Fig. 1). Paradoxically, London has one of the most developed landscape protection systems. A comprehensive system of assessing and selecting city views was developed supported with detailed description of advantages of the existing landscape which needs to be protected [26]. This is accompanied by a system of verifying planned tall buildings by simulating their impact in selected strategic views: London Panoramas, River Prospects, Townscape Views and Linear Views [16]. However, the speed of spatial changes exceeds any controlled levels. It is not easy to counteract planning errors made several decades ago, and frequently those errors support consecutive cases of mismatch in planning locations for tall buildings. In the 20th c., when the first tall buildings were built, the tool kit of an urban planner was limited to traditional tools, such as sketches and individual cross sections. Professionals could resort to their experience and intuition only. Thus, a number of design errors were made. Is the contemporary methodology of analysing locations of tall buildings sufficient for determining relevant and comprehensive spatial policies?

The majority of cities which faced the issue of tall buildings developed modern instruments for monitoring their impact on the landscape. Such instruments for London [16, 26] are some of the most complex examples. They include detailed information concerning major views of the city, criteria for selection and significance

of its cultural heritage. Each new construction project, pertaining to developing a tall building in particular, necessitates analysing the impact of a building facility on the landscape within predefined strategic views [16]. A similar procedure is applied in Koln [14] and Dusseldorf [13]. Several cities have also developed strategies for locating tall buildings [1, 4, 22, 28]. The strategies usually comprise studies of urban morphology, which examine conditions at specific sites for developing tall buildings. They define potential locations while taking into consideration factors related to the overall composition (e.g. shadow effect, accessibility, public transport, land topography, air traffic, etc.). Interesting studies have been developed in Ottawa (2007) aimed at protecting the landscape of the Parliament Hill [5] and in Vancouver (2011) to determine the potential for developing the silhouette of the city and encompassing a mountainous landscape in the background [15]. The study developed for Milan (2007-2011) focused on analysing visual experience while moving within the structure of the city. The analysis was supported by a 1:500 mock-up of the city and a special micro-camera used for simulating the reality [21]. A study developed for Warsaw (2008) is more general in its nature. It defines tendencies for developing tall buildings based on the silhouette of the city seen from the side of the river [19]. According examples of tall buildings developed recently in the city, so much limited criteria turned out to be insufficient to protect valuable urban and architectural facilities. This short overview of techniques for analysing influence of tall buildings on the landscape show how diverse approaches used can be and highlights the fact that there is no universal standard in the area. We are still searching for relevant solutions. The use of digital tools, however, is a step towards obtaining objective results and expediting analyses.



Figure 1: Axial view from Millennium Bridge, London, with tall buildings seen behind St. Peter's Cathedral

2. VIRTUAL CITY MODELS AS NEW TOOL FOR URBAN ANALYSIS

In recent years, the number of 3D virtual models and their precision were growing at an exponential rate. Texturing of elevations and integrating models with photographs (Google Street View) provides much higher degree of virtual imaging the reality. A pre-taste of a further advancement is a new standard of data recording, CityGML [11], developed recently. While analysing the development trends, we may assume that probably within coming years 3D models will cover entire territories of economically developed countries, fully reflecting their built-up substance and natural landscape. But, for the urban planning virtual city, models are only a new tool. Neither the degree of reality nor interactivity of visualization is important. What counts is the possibility of application for analyses and simulations, extending the conventional, manual urban planning workshop [23].

To assess the impact of tall buildings, it is necessary to visualize panoramas [6]. This is a specific form of projecting city space combining various development planes overlapping

each other. Numerous experiments showed that perception of such views depends much on facilities with distinguishing building which stand out in their surroundings: contemporary and historic landmarks, existing skyscrapers of unique silhouettes, church towers, characteristic historical buildings, some elements of infrastructure and other visually accentuated elements [23]. On the one hand, it is important to differentiate model precision levels as required by specific analyses. Certain elements can be presented at a general level others need to be extremely precise. On the other hand, in the majority of planning applications, 3D city models need to be linked with a relevant database (e.g. functions of facilities) and support interpretation of structural relations (e.g. recognition of building as a whole rather than set of walls or lines).

Research on using virtual city models in urban planning has been rapidly growing. However, the potential of the tool is still very little recognised. The development of tools for developing and imaging of those models is ahead of their application. The main stream of application of virtual city models is still focused on presentation and visualization of city structure for facility services, commercial sector and marketing, promotion and learning of information on cities (like: promoting cities, tourism, visualizations for investors, etc.) [20].

2D and 3D GIS analyses are considered a more advanced way of using city models. Readymade and optimized tools dedicated to the Geographic Information System can be used for: measuring density of buildings, studying proximity, spread analysis, like dynamic air flow, acoustic analysis, simulation of disasters, traffic management etc. [17]. For instance, 3D models are used in visibility analyses. The visibility of points in a landscape from one or more locations has many applications. These include studies of scenic quality, urban design, civil and military observation needs and telecommunications planning, amongst others. Typically these operations are performed on grid files, but a similar, vector-based procedure

is known as Isovist analysis [10]. The relevant theory in this field is still under development. Consequently, fields of application also increase in numbers. The development of the theory has been examined by Fisher-Gewirtzman [12] while referring to the most important publications. Among them, a particular classical role is played by Benedikt [3] who was the first to introduce the Isovist and to develop a set of analytic measurements of isovist properties. A number of researchers have developed tools for automated isovist analysis. Turner et al [29] uses a set of isovist to generate a graph of mutual visibility between locations. Batty [2] describes how a feasible computational scheme can be used for measuring isovist fields and illustrated how they can visualize their spatial and statistical properties by using maps and frequency distributions. Yin [30] in his doctoral thesis summaries limitations of 2D and 3D visibility calculations.

We could also witness a major progress in developing tools supporting calculations. Simple and intuitive applications for obtaining a map/visualization of the viewshed for a particular point in space were in recent years developed. Google Earth Pro provides a tool to visualize the field of visibility in the 3D models available on the platform [31]. The precision of the analysis, however, is still low and cannot be treated as a research tool. GIS tools provide seem to be more promising, for example ArcGIS by Esri. They enable determining the visibility of sight lines, identification of observer points from raster surface or construction of sight lines, among others. [32] The number of advanced GIS tools is growing. The same applies to GIS open source software. The functionality of free QGIS together with GRASS GIS package is similar to Esri. The number of possible field of application of those tools has been growing as well. One of them is diagnosis and simulation of tall buildings.

3. SKYLINE AND TALL BUILDINGS – SELECTED METHODS OF ANALYSIS

This chapter discusses selected methods used by the author of the article that are crucial for examining geometrical issues determining the influence of tall buildings on the skyline. The theoretical framework of those methods was initially developed between 2004–2006 and described in a doctoral thesis [7]. Later, the methods were further developed (since 2012, in Cyber Urban Centre at West Pomeranian University of Szczecin, Poland). Relevant software was developed supporting the use of new tools in urban planning.

3.1 Analysis of visual impact range

Planning of new tall buildings necessitates analysing the urban structure of a city at various scales: from global, including the impact of a building on the space of the entire city, partial external exposition within skylines, to internal views of public space (squares and streets) [31]. Visual perception of a city is a dynamic process. Relations between buildings change together with the point of observation. These relations are analysed against lines of buildings and visual planes. A photography of a skyline shows a part of the impression only (limited to one point in space). For planning purposes, a relevant synthesis is necessary – determining the sum of visibility fields for a planned tall building.

The method aims at providing objective imaging as regards range and visual impact of a single tall building. The algorithm is based on a geometrical analysis of a virtual city model. The only requirement is to specify a precise location of a tall building. Computer simulation produces a map with all locations from which the planned building can be seen. The impact area can be presented in a projection and in two axonometric or perspective views. Usually, the result is limited to examining of public space, understood as all undeveloped sites in a city. Simulation, however, may include examining of tall building visibility from other buildings and all predefined geometrical elements of the

3D city model (Fig. 2).

Objectives of the method are similar to those of isovists and methods of automated determination of visibility fields available in GIS tools. Differences apply to the interpretation and precision of simulations. A novelty of the method is imaging of not only real visual impact range but also imaging of the impact power (expressed in intensity of colour used). Complexity of results depends on the complexity of the city structure. It obviously depends on the precision of the virtual city model, including the precision of the site digital model. Taking into consideration of tall green (various ways of modelling) is important, since it may have its impact on the result of the simulation. Results of simulations using the method frequently deviate from our intuitive expectations. Geometrically defined impact of a tall building is usually much broader than we could assume and the computer analysis reveals new and not that obvious spatial relations in the city.

3.2 View angles analysis

Another direction for developing urban analysis is to examine particular viewpoints. In the case of several studies, protecting specific strategic views is an important element of the landscape development strategy. The study developed for London [16] lists 27 views which are strictly protected and are considered to be very important from the cultural heritage point of view. Traditional urban planning uses specific terms to describe a morphology of a view, including view foreground and exposition background. However, it is quite difficult to define geometrically and measure mutual relations between various planes of a view. Every change in the position of an observer changes the configuration of a view and relations between particular elements of the exposition. Research should cover all variables, including land topography, distance between buildings, building height, tall green, distance and height of an observation point. Due to the complexity of spatial relations between various buildings in the city it is necessary to find techniques for

assessing the significance of particular architectural facilities for the overall composition of the city skyline.

The method of analysing view angles aims at an objective assessment of the impact of particular buildings on a given observation point. The analysis is based on calculating an angle between the observer's eye and the tallest point of each building. The largest angle is in the case of tall buildings and those situated close to the observation point (Fig. 3). By using classical designing techniques, usually involving projections and cross sections of the city, it is possible to analyse relations between selected buildings only. Virtual spatial models enable us to expedite and automate procedures, as well as significantly increase the range of buildings analysed to cover all facilities in the spatial model. Analyses take into consideration the exposition height and for each building: its height above sea level, relative height above ground level and distance to the point of exposition (Fig.4).

Thus, the simulation produces view angle maps. The algorithm used enables diverse imaging. The term of a building can be interpreted in various ways (frequently applied division into smaller elements). Contrary to traditional techniques, results are visualised in cross sections but in projections, including axonometric ones. A 10 degree scale is the most frequently used in studies (covering various angle ranges). Although, objectives of the method are relatively simple and based on elementary geometrical operations, the method itself is not popular. On the one hand, there is a shortage of readily available tools in CAD/GIS. This may result from the fact that the simulation results in a quite abstract image of a city, which does not take into consideration a fact that buildings obscure each other. On the other hand, in many instances, such a general image can be very valuable for assessing the interpretation of a view and distinguishing specific exposition planes.

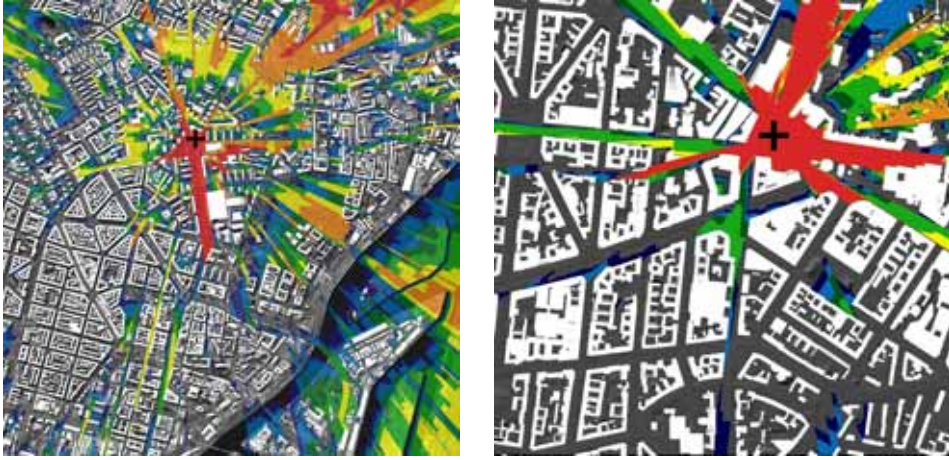


Figure 2: Analysis of visual impact range of tall building; Hanza Tower in Szczecin, Poland (left), Zoofenster Berlin (right)



Figure 3: Illustration of the idea of view angles analysis (London)

3.3 The visibility coverage analysis of panoramic view

The method of view angle analysis provides for emulating general impact maps from a given observation point. While analysing city landscapes and studying tall buildings it is important to determine a visibility field taking into consideration mutual impact of buildings in terms of obscuring their view. Skylines expose certain set of elements of the city structure. Some buildings are more visible than others. A view comprised also trees, bridges, technical infrastructure and, first and foremost, topography (e.g. hill in background or foreground). All these elements comprise a panoramic view limited with a skyline of the city. It is possible to determine the panoramic view as a 2D image.

However, all elements of the image have their third dimension and they are more or less dispersed in the real space of a city.

The method aims at determining which elements comprise a given view and imaging of those elements together with their spatial relations. While analysing tall buildings, we focus most often on panoramic views. However, any view of a square or a street can also be examined. The basis for the application of the algorithm is a virtual city model. Coordinates of an observation point are pre-set. And the result is a spatial set of elements seen from the observation point (e.g. parts of walls and sections of land, etc.). The analysis is a kind of spatial interpretation of the visible planar projection (Fig. 5). Precision depends of accuracy of the 3D city model and calculation power of a

computer. Objectives of the method are similar to those of isovist described by Benedikt [3]. Isovist applications in urban planning and architecture are limited, however, to 2D studies. Determining of visibility in GIS software is frequently limited to a digital model of a given area. Yet another difference is the accuracy of results.

To be able to study the impact of tall buildings on the landscape, it is necessary to use various analytical methods. Selection of those methods depends on accuracy and recording of the model, as well as specific nature of a city landscape (topography, dispersion of urban structure, distribution of historical and contemporary dominants, etc.). For example, the combination of the view range method presented and analysis of view angles can produce interesting results. Examples of such simulations are presented in figure 6. The combination of the two methods enables determining of a build-up zone obscured by foreground buildings and obtaining information about angular values. Contrary to an open landscape, the skyline observed is distributed in space and divided into pieces. Information about angles for visible buildings may help determining height (and thus angular value) for new buildings in a given area. Results are important while determining rules for visual protection of important urban development. Planned buildings within a skyline should not have higher angular values from those in the foreground.

4. APPLICATION OF PRESENTED METHODS IN PLANNING

The methods discussed in the previous chapter supporting planning and analysing the impact of tall buildings on a city landscape and city skyline were applied in several urban studies in 2005-2011. Interaction between theory and practice, initiated in the first phase of developing those methods, was crucial while formulating their objectives. The studies were implemented by the Cyber Urban Center (West

Pomeranian University of Technology of Szczecin, Poland) involving the author of the article.

4.1 Studies of several tall buildings in Szczecin, Poland

One of examples of applying analytical methods and computer aided tools were studies of tall buildings impact on landscape in selected locations in Szczecin in 2007 (Fig. 7). The studies were implemented under a contract with the local government [9]. The time for developing those studies was special. Immediately before the first wave of the global crisis (in 2008), strategic investors in Europe showed increased interest in developing tall buildings. The trend was very clear in developing European countries, such as Poland. At that time, in Szczecin, plans were developed to build several tall buildings. The studies analysed in total 10 potential investment projects. The aim was to determine the impact of planned facilities on the city landscape while taking into consideration cultural values and define detailed guidelines concerning their height and form.

The analysis was based on a virtual model of Szczecin developed by the CUC in 2005-2007. While implementing the objective, it turned out that the visual impact range method was particularly useful (Chapter 3.1). The algorithm of the method was extended and optimized several times. Finally, a series of impact maps was produced for each location. Complexity of results exceeded any expectations. One some of them were in line with intuitive assumptions.

The group of facilities analysed included the tallest of buildings planned Hansa Tower (125 m). The analysis of its visual impact showed directions of exposition along axis of major streets and enables determining an optimized position of the facility at its plot. It also enabled setting a field of exposition from the river and assessing the relationship between the new building and the historical skyline of the city. Skyline simulations were developed for

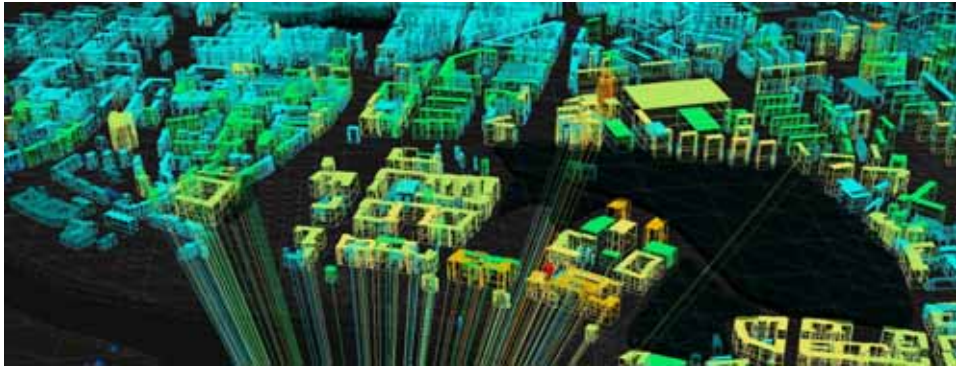


Figure 4: View angles analysis for panoramic view, Szczecin (Poland)



Figure 5: Visibility range analysis of panoramic view, Szczecin (Poland)

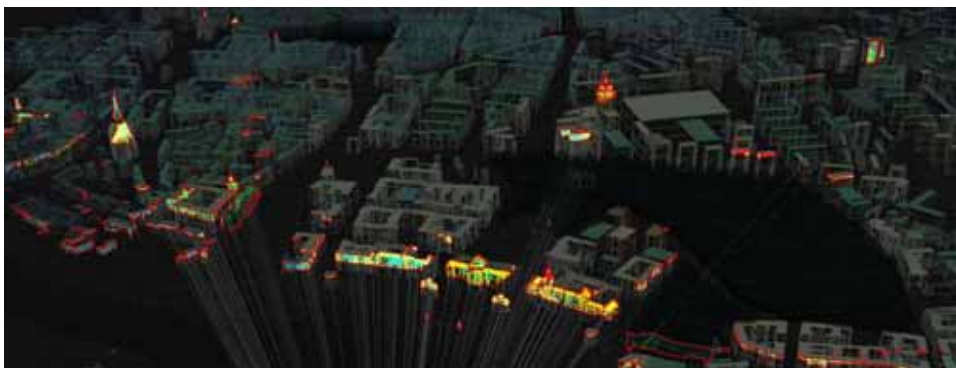


Figure 6: Combination of the view range method and analysis of view angles, Szczecin (Poland)

areas of the strongest impact (Fig. 2). The studies helped excluding some locations for developing tall buildings, and in other locations such facilities were allowed provided minimum height was preserved. Guidelines became a part of master plans which comprise local law. Until 2013, about 30% of plots were developed according to those guidelines.

4.2 Protection of historical panorama of Lublin, Poland

Yet another example of the application of methods (presented in Chapter 3) is the 'Study of visual values of the city of Lublin' developed under a contract with the local government in 2011 [8]. It aimed at determining rules for protecting historical skylines and silhouettes. Basic questions for the project team included: What should be the direction for the development of the city to preserve panoramas of the invaluable in Europe old town of Lublin? How can their exposition be strengthened? Is it possible to erect tall buildings in Lublin?

The tasks mentioned above and specific city landscape were decisive as regards selecting tools for their implementation. A limiting factor was input material important for further digital analyses. The city did not have its 3D model ready and there was shortage of time and finance for developing it. Therefore, it was not possible to use computer methods. The only input available was a digital model of the area. Based on 2D land register maps, the team developed analytical models of buildings [24]. While determining zones for protecting views of the old town, two methods described in Chapter 3: analysis of view angles and analysis of visibility range turned out to be particularly useful (Fig. 8).

Determining of the exposition background zone was an important task under the study. It was very difficult to define it geometrically. Boundaries of the zone were determined by combining the two methods. Synthesis of findings from the two methods enabled establishing areas particularly sensitive to new investment. The team observed major relations between

angular values and topography, and thus visibility of buildings in space. Findings of the study were included in the current strategy for spatial development of Lublin.

5. CONCLUSIONS

Tall buildings in European cities are usually give rise to various controversies, discussions and disputes. On the one hand, we have trends to preserve cultural values of a historically developed landscape and skylines, whereas on the other arguments are raised that a modern image of a city should be created with tall buildings considered as a reflection of advancement, modernity and prestige. Wrong decisions are sometimes based on insufficient knowledge about spatial consequences of buildings planned. For this reason, objective findings of geometrical methods used for landscape analysis is crucial. Another reason for promoting scientific research on city landscape is a rapid development of geo-information techniques, which results in growing availability and precision of virtual city models. It is a new tool in urban planning and the potential for application is not fully explored.

Analytical methods presented in the article provide for a broader assessment of the impact tall buildings have on a city landscape if compared with more traditional techniques used in urban planning. Geometrical analysis is a basis for algorithms used. By its nature, results are objective and cannot be debatable. Doubts may apply to accuracy of simulations and relevance of the methods as regards the implementation of specific planning goals. According to previous experience in applying the methods, two conclusions can be drawn. Firstly, visual impact of tall buildings on the city landscape is a complex issue. Even in medium sized cities, such as Szczecin and Lublin, majority of spatial interaction can hardly be foreseen intuitively. Secondly, the interpretation of findings is an equally complex and tedious process. Imaging of results is frequently quite abstract. Similarly to the interpretation of a magnetic resonance image in medicine, the analysis of results de-

rived from methods concerned require considerable competence in using the tool and urban knowledge.

The research methods have their limitations. Quality of results depends on the accuracy of a 3D city model used and its format. Not all models can be used for analytical analysis. An important obstacle is the power of computers and effectiveness of algorithms used in applications. Further research focus on optimizing and developing new methods of urban simulations using the most recent computer tech-

niques, such as simulating impacts or tall buildings on complex spatial structure of cities, delimitating areas of protection of city panoramas and historically originated urban interiors and introducing multi-aspect systems of urban analyses (3D-UAS) to combine spheres of geo-information with urban planning, which is the main focus of the 2TALL Project (Application of 3D Virtual City Models in Urban Analyses of Tall Buildings) currently under implementation by the author and the CUC team.



Figure 7: Simulation of planned tall buildings in city skyline (Szczecin, Poland 2007)

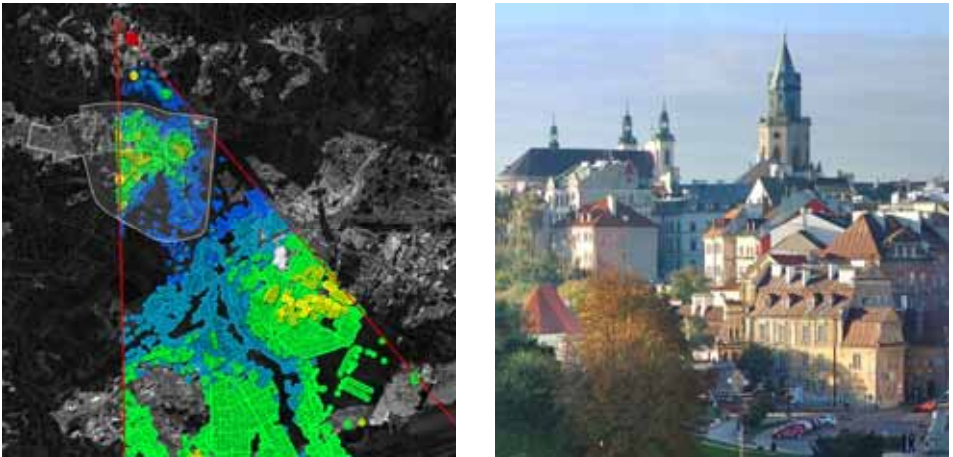


Figure 8: A fragment of historical panorama of Lublin (Poland) and computer analysis with view angle method

ACKNOWLEDGMENTS

This research was funded by a Norwegian Financing Mechanism. Digital model of Berlin provided by Berlin Partner GmbH. All other 3D city models used for simulation were made by Cyber Urban Center at WPUT Szczecin. I gratefully acknowledge this support.

REFERENCES

- [1] Bankside, Borough and London Bridge Opportunity Area : Stage 1 - Tall Building Research Paper, Study for Southwark Council, London 2010.
- [2] M. Batty, Exploring isovist fields: space and shape in architectural and urban morphology, *Environment and Planning B: Planning and Design*, volume 28, London 2001, pp.123-150.
- [3] M. L. Benedikt, To take hold of space: isovist fields, *Environment and Planning B: Planning and Design*, vol. 6, London 1979, pp. 47-65.
- [4] Brighton & Hove Tall Building Study, Study for Brighton & Hove City Council, by Gillespies and GVA Grimley, Brighthelm 2003.
- [5] Canada's Capital Views Protection. Protecting the Visual Integrity and Symbolic Primacy of Our National Symbols, collective work for National Capital Commission, Ottawa 2007.
- [6] K. Czyńska, Using a model of virtual city for research on visibility range of panoramas of the city, *Space and Form* 2009 no 12, Szczecin, pp. 111-114, ISSN 1895-3247.
- [7] K. Czyńska, Metody kształtowania współczesnej sylwety miasta na przykładzie panoram Szczecina – wykorzystanie wirtualnych modeli miast w monitoringu i symulacji panoram. Doctoral thesis, Wrocław University of Technology, Wrocław 2007.
- [8] K. Czyńska, W. Marzęcki, P. Rubinowicz, Wartości widokowe miasta Lublin, Urban Study commissioned by Department of Spatial Planning, City of Lublin, Lublin 2011.
- [9] K. Czyńska, W. Marzęcki, P. Rubinowicz, Analyses of visual impact and definition of spatial guidelines for high buildings in Szczecin, Urban Study commissioned by Department of Spatial Planning, City of Szczecin 2007.
- [10] M. De Smith, M.F. Goodchild, P.A. Longley, *Geospatial Analysis. A Comprehensive Guide to Principles, Techniques and Software Tools*. 3rd edition. Viewsheds and RF propagation 2009, on-line version <http://www.spatialanalysisonline.com>
- [11] S. Donkers, Automatic generation of CityGML LoD3 building models from IFC models, Msc thesis, Delft University of Technology 2013.
- [12] D. Fisher-Gewirtzman, 3D models as a platform for urban analysis and studies on human perception of space. Usage, Usability, and Utility of 3D City Models 01001 (2012), *EDP Sciences*, 2012, DOI: 10.1051/3u3d/201201001.
- [13] Hochhausentwicklung in Düsseldorf Rahmenplan. Beiträge zur Stadtplanung und Stadtentwicklung in Düsseldorf, collective work under the direction of Richard Erben, Düsseldorf 2004.
- [14] Hochhauskonzept Köln 2003, Stadtplanungsamt der Stadt Köln, 2003.
- [15] Implementation of "Vancouver Views" and Opportunities for Higher Buildings in the Downtown, study for Standing Committee on Planning and Environment, Vancouver 2011.
- [16] London View Management Framework. Supplementary planning guidance, Study by Greater London Authority, Mayor of London 2012.
- [17] J. Moser, F. Albrecht, B. Kosar, Beyond visualisation – 3D GIS analyses for virtual city models. *ISPRS 5th International 3D*

- GeoInfo Conference, Berlin 2010: International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, pp. 143-147.
- [18] R. Musiał, Supertall Buildings – Present and Future Components of the Landscape of Big Cities in Europe, Technical Transactions, Architecture, Issue 1-A/1/2012, Krakow 2012.
- [19] W. Oleński, Digital City Panorama. GIS in analysis of the landscape of Warsaw, Arcana GIS, Warszawa 2012, pp. 27-30.
- [20] S. Pal Singh, K. Jain, V. R. Mandla, Virtual 3D city modeling: techniques and applications. ISPRS 8th 3DGeoInfo Conference, Volume XL-2/W2. Istanbul 2013: International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, pp. 73-91.
- [21] B. Piga, V. Signorelli, E. Morello, Anticipating the Impacts of Urban Design Projects Starting from the Pedestrians' Experience, AESOP 26th Annual Congress METU, Ankara 2012.
- [22] Plymouth Tall Buildings Strategy (draft), Plymouth City Council, Plymouth 2005.
- [23] P. Rubinowicz, Cyber Urban Center: The Visual Impact Simulations for Tall Buildings Analyzes in Szczecin, Creative Urbanism, National University Lviv Polytechnic, Lviv 2014, in press.
- [24] P. Rubinowicz, Various Aspect of Urban Structure Analysis while Assessing the Friendliness of a Place – Example of Lublin, Urban Landscape Renewal no 6, vol. 2, Gliwice 2012, pp. 345-350.
- [25] M. Schwarzer, Architecture and Mass Tourism, [in:] J. Ockman, S. Frausto (edit.) Architourism, New York 2005, pp. 24-25.
- [26] Seeing the history in the view. A method for assessing heritage significance within views, Study by Historic Buildings and Monuments Commission for England, London 2011.
- [27] Suleiman Wassim, Joliveau Thiery, Favier Eric (2013) A New Algorithm for 3D Isovist, in: Geospatial Dynamics, Geosimulation and Exploratory Visualization, Springer Berlin Heidelberg 2013, pp 157-173
- [28] Tall Buildings Strategy Swansea, City and Country Council of Swansea, White Young Green, Swansea 2008.
- [29] A. Turner, Analysing the visual dynamics of spatial morphology, Environment and Planning B: Planning and Design Vol. 30, London 2003, pp. 657-676.
- [30] J. Yin, Mobile 2D and 3D Spatial Query Techniques for the Geospatial Web, Doctoral Thesis, Dublin Institute of Technology, 2013.
- [31] A. Zwoliński, Analysis and parameterization of public spaces using 3D city models, National University Lviv Polytechnic, Lviv 2014, in press.
- [32] <https://support.google.com/earth/answer/3064261?hl=en>
- [33] <http://en.wikipedia.org/wiki/ArcGIS>

ABOUT THE AUTHORS

Klara Czyńska, PhD architect, assistant professor at West Pomeranian University of Technology Szczecin, Poland. PhD work about application of virtual modeling of cities for analysis of high-rise buildings. She has also participated in planning studies in this field. Currently she is a leader of 2TaLL research project (*Application of 3D Virtual City Models in Urban Analyses of Tall Buildings*) funded by the Norwegian Financial Mechanism. Since 2013, member of Cyber Urban Center (CUC) research team. Contact: kczyńska@zut.edu.pl

GEOMETRY AND FUZZY NAVIGATION SYSTEM FOR VIRTUAL ROBOT IN THE DRAWING ENVIRONMENT

Algirdas SOKAS

Vilnius Gediminas Technical University, Lithuania

ABSTRACT: This article analyzes robot orientation problem in the unknown environment. This is an idealize task that a virtual robot has to solve seeking to find its way in the drawing with graphical objects obstacles. The virtual robot uses the five graphical sensors system. The fuzzy navigation system control three states target, obstacle and workaround. All states present robot's steering angle. Integrated CAD environment (AutoCAD) with Visual Basic Application programming language enlarges possibilities in modeling fuzzy navigation systems. Fuzzy system, examples of virtual robot presented and conclusions made.

Keywords: Fuzzy system, navigation, virtual robotics.

1. INTRODUCTION

I am very interested in artificial intelligence. I wanted to understand how a computer program can make decisions on how a robot can find a free route in an unknown environment and how having several parameters it makes the right decision.

The basic idea of underlying fuzzy logic control was suggested by Lotfi A. Zadeh [12-14]. A fuzzy linguistic label can be represented by a fuzzy number which is represented by a fuzzy set. Fuzzy sets capture the ability to handle uncertainty by approximate methods. A triangular fuzzy number applied mostly in the fuzzy theories and applications.

The first implementation presented by Mamdani and Assilian [7] in connection with the regulation of a steam engine. In the ensuing years, once the basic idea underlying fuzzy logic control became well understood, many applications followed [4, 8]. In particular, fuzzy logic-based approaches are successful in dealing with complex systems whose mathematical models are not available, where the conventional model-based approaches would fail. Fuzzy logic attempts to apply a human-like way of thinking in the application areas. Fuzzy

logic is basically a multivalued logic that allows linguistic terms for intermediate values to be defined besides conventional evaluations. Fuzzy control system employ a mode of approximate reasoning that resembles the decision making process of humans. The behavior of a fuzzy controller, can easily understood by a human expert as the knowledge expressed by means of intuitive, linguistic rules. A fuzzy system usually designed by interviewing an expert and formulating his implicit knowledge of the underlying process into a set of linguistic variables and fuzzy rules [11].

Mobile robot local path planning in an unknown environment with uncertainties is one of the most challenging problems in robotics. For autonomous navigation, the robot should be capable of sensing its environment, interpreting the sensed information to obtain the knowledge of its position and the environment, planning a route from an initial position to a target with obstacle avoidance, and controlling the robot direction to reach the target [9]. In mobile robot computer program the navigation procedure is applied iteratively until the robot reaches its final destination [2]. The fuzzy logic control applied to mobile robot navigation and obstacle avoidance has investigated by several research-

ers. The control based on this theory with virtual robots [5] provides satisfying results. Many application works of fuzzy logic in the mobile robot have given promising results [1, 3, 6], comparison of the related methods that address to mobile robot navigation problem presented in [10].

In the article for decision making logic part used Mamdani inference method and for de-fuzzification is used center of area method (COA) [4]. AutoCAD is a program used as operating environment, and Visual Basic for Application (VBA) is a language used for programming. Author presented virtual robot working in the drawing environment.

2. A FUZZY NAVIGATION SYSTEM

In the drawing we draw a rectangle which we call the robot as it, using its fuzzy navigation system, will bypass obstacles and move to the target object. Robot step is from point a to point b (Figure 1). It is simply copied from point a and moved to point b with the angle α turn presented by the system. Robot-to-target distance D is known as the program states points a and t coordinates.

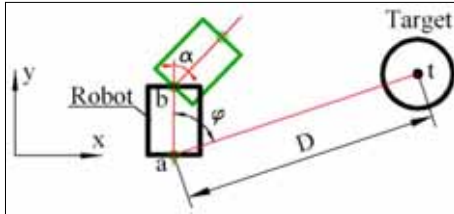


Figure 1: The virtual robot moving system and target object.

A drawing is a very interesting environment in which we can simulate the real environment. The robot has graphical sensors. One is the front sensor that mimics a line drawn through the robot points a, b and extending 30 mm; if the intersection point c with obstacles is found, we have the front sensor value $F = bc$ (Figure 2).

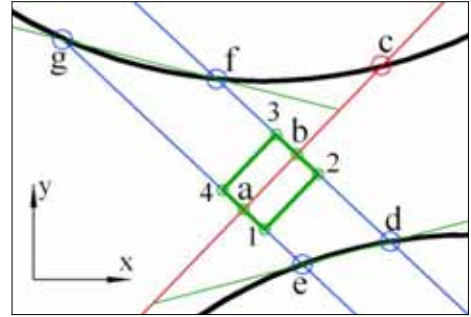


Figure 2: The virtual robot with sensors system and coordinates.

On its right are two sensors, one of which is produced by a line from point b to point 2 and extending 30 mm; if the intersection point d with obstacles is found, we have right side sensor value $Sr = 2d$ (Figure 2). The other right sensor is equal to $Sr2 = 1e$ distance. Analogously the left sensors distances are found $Sl = 3f$ and $Sl2 = 4g$.

This can be done programmatically. A line extension performed with SendCommand "lengthen" operation, and an intersection with an obstacle captured with object.IntersectWith operation.

We draw a lines between intersection points e, d and g, f to the robot axis passing through a, b and find angles β_r, β_l of the tangents, which show the object's approach angles to the obstacles (Figure 2).

Hence, we have a whole range of parameters and must decide what angle to rotate the robot in the next step. We will use the fuzzy navigation system, which has three states: target, obstacle, and workarounds. The target state controlled according to the distance to target D and angle ϕ determining steering angle. The obstacle state is controlled according to the front sensor F and side sensor Sr or Sl values determining steering angle. The Workaround state is controlled according to the angle β of the tangents and side sensor Sr or Sl values determining steering angle.

We are dealing with the problem of obtaining fuzzy navigation system. We will analyze fuzzy rules in the next chapter.

3. FUZZY RULES

In the first state the robot moves towards the target. Distance D to the target captured and presented in four membership fuzzy function (Figure 3). Robot's orientation φ with respect to target presented in seven membership fuzzy function (Figure 4).

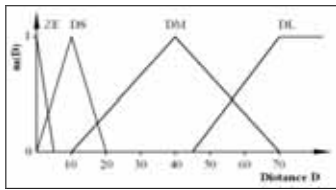


Figure 3: Distance to target D , there ZE – zero, DS – small, DM – medium, DL – large.

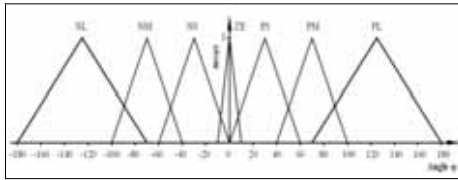


Figure 4: Robot orientation angle φ , there NL – negative large, NM – negative medium, NS – negative small, ZE – zero, PS – positive small, PM – positive medium, PL – positive large.

Robot's steering angle α , which is presented in seven membership fuzzy function (Figure 5), is determined based on its distance D and orientation angle φ . The fuzzy rule matrix for steering angle (Table 1) made according to these three graphs, whose columns named after distance memberships and rows named after orientation angle memberships, and the matrix filled with the appropriate robot steering angle values. This matrix allows to programmatically automate the entire fuzzifica-

tion process.

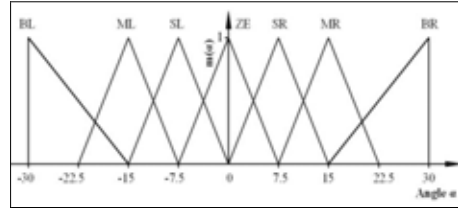


Figure 5: Robot steering angle α , there BL – big left, LM – medium left, SL – small left, ZE – zero, SR – small right, MR – medium right, BR – big right.

Table 1: Robot steering angle α dependent from D and φ .

		D			
		ZE	DS	DM	DL
φ	NL	BL	ML	ML	SL
	NM	BL	BL	ML	ML
	NS	BL	ML	ML	SL
	ZE	ZE	ZE	ZE	ZE
	PS	SR	MR	MR	BR
	PM	MR	MR	BR	BR
	PL	SR	MR	MR	BR

In the second state the robot detect the obstacle. Side distance S to the obstacle captured and presented in fifth membership fuzzy function (Figure 6). Robot's front distance F with respect to obstacle presented in third membership fuzzy function (Figure 7).

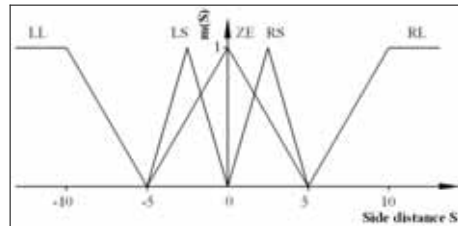


Figure 6: Robot side distance S , there LL – left large, LS – left small, ZE – zero, RS – right small, RL – right large.

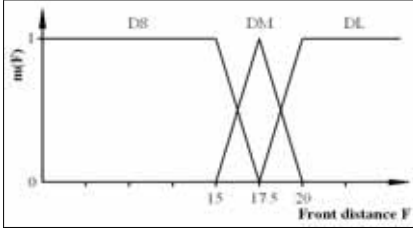


Figure 7: Robot front distance F , there DS – distance small, DM – distance medium, DL – distance large.

Robot's steering angle α , which is presented in seven membership fuzzy function (Figure 5), is determined based on the its front distance F and side distance S . The fuzzy rule matrix for steering angle (Table 2) made according to these three graphs, whose columns named after side distance memberships and rows named after front distance memberships, and the matrix filled with the appropriate robot steering angle values.

Table 2: Robot steering angle α dependent from S and F .

		S				
		LL	LS	ZE	RS	RL
F	DS	BR	BR	BL	BL	BL
	DM	BR	MR	ML	ML	ML
	DL	MR	SR	SL	ML	ML

In the third state the robot workarounds near obstacle. Tangent angle β to the obstacle captured and presented in fifth membership fuzzy function (Figure 8). Robot's side distance S with respect to obstacle presented in fifth membership fuzzy function (Figure 9).

Robot's steering angle α , which is presented in seven membership fuzzy function (Figure 5), is determined based on the its tangent angle β and side distance S . The fuzzy rule matrix for steering angle (Table 3) made according to these three graphs, whose columns named after tangent angle memberships and rows named after side distance memberships, and the matrix filled with the appropriate robot steering angle

values.

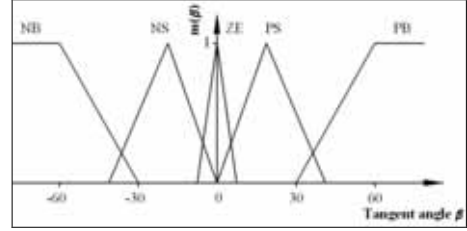


Figure 8: Robot tangent angle β , there NB – negative big, NS – negative small, ZE – zero, PS – positive small, PB – positive big.

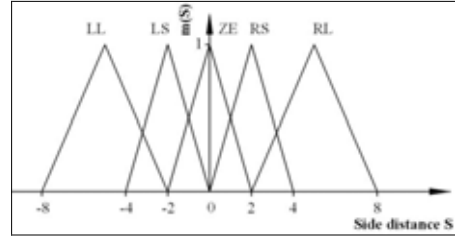


Figure 9: Robot side distance S , there LL – left large, LS – left small, ZE – zero, RS – right small, RL – right large.

Table 3: Robot steering angle α dependent from β and S .

		β				
		NB	NS	ZE	PS	PB
S	LL	BR	MR	SR	ML	BL
	LS	MR	SR	ZE	SL	ML
	ZE	SR	SR	ZE	SL	SL
	RS	ML	SL	ZE	SR	MR
	RL	BL	MR	SR	SR	BR

So we have three states: target, obstacle and workarounds. All situations present robot's steering angle α . An application is required that manages these states depending on the information received from the graphical sensors.

4. ALGORITHM AND PROGRAM

At the beginning the virtual robot is at the start point and turned by the original angle. The program emulates robot and cyclically rotates by the specified angle. Program's algorithm presented in Figure 10. The robot turns according to the procedure Target towards the target and moves towards it when it finds a straight line. If an obstacle is met the Obstacle procedure is called and the robot turns to the left. In the event of an object near a barrier it captures the right side distances and tangent angles. Then steering angle set by the Workarounds procedure. If the sensor captures the left side distance and the left tangent then the Workarounds procedure operates on this data. If the road is clear, the robot moves the shortest distance to its destination.



Figure 10: Algorithm of working three states: target, obstacle and tangent.

Algorithm is presented by a program consisting of four classes which manage the robot and its three situations. Software sensors controlled by five module procedures (front distance, left and right side distances, left and right side tangents). The relationship between the graph and the matrix ensured by three more module procedures because there are three types of graphs.

5. EXAMPLES

Here are some examples of conduct of the vir-

tual robot in the graphical environment. The robot moves towards the goal and is bumps into a circle (Figure 11). In the fifth step distance $D < 30$ mm runs the Obstacle procedure, the robot turns left (based on the algorithm), and in the next step it turns to the left even more. The lateral sensors show the values and when they become $S < 30$ mm it starts the third procedure Workarounds. The robot starts to move near the obstacle. When the shortest distance to the target object is exposed, the robot turns towards it.

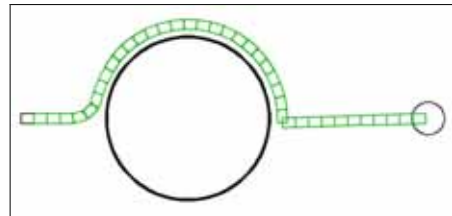


Figure 11: An example of the virtual robot motion in the drawing with circle.

How virtual robot acts between several circles on the way towards the target object (Figure 12). At first the robot avoids the first circle until the left lateral sensor shows lower values than the right and the robot starts to bypass the second circle (Figure 13), but the way is free and a few steps straightens towards the target. The robot encountered a third circle (Figure 14) and it changes direction.

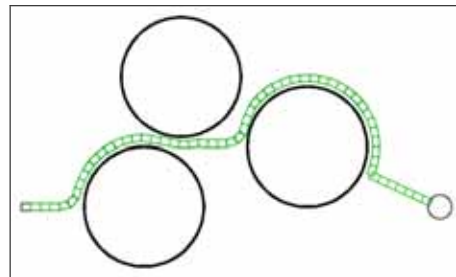


Figure 12: An example of the virtual robot motion in the drawing with circles.

The figure (Figure 13) shows a fragment

with only lateral graphical sensor operating positions. Intersecting lines and circles points are noted, distances are calculated and tangents are determined. The figure (Figure 14) shows a fragment with only the front graphical sensor operating positions. Intersecting lines and circles points are shows as small circles, distances are calculated and decisions are made.

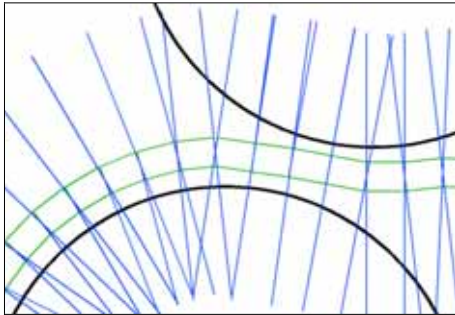


Figure 13: The lateral graphical sensor operating positions.

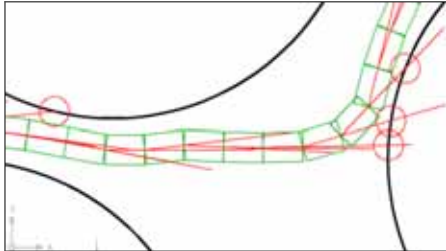


Figure 14: The front graphical sensor operating positions.

How the robot acts when it meets a more advanced obstacle formed of the lines shown in Figure 15. At the beginning the robot is moving toward a goal based on the first state, when an obstacle is met, it turns to the side and looks at a few of equations towards the goal, but the obstacle extends to the opposite direction and the robot avoids the whole obstacle until a free path towards the target is clear.

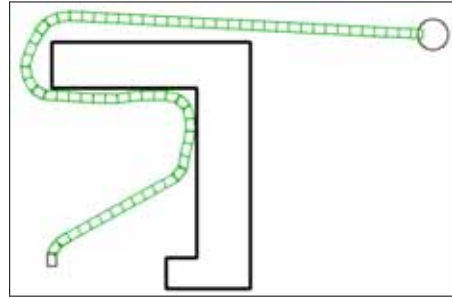


Figure 15: An example of the virtual robot motion in the drawing with lines.

In this example, the graph exhibits high sensitivity. The robot behaves as if intellectual.

6. CONCLUSIONS

A utility is designed that allows to go deeper into the decision made by a computer program how a robot can find a free route in an unknown environment having several parameters to make the right decision.

By merely programming I finally learned how to assess the two charts with four parameters and obtain only one value. The diagrams are bound by matrices and functions provide a definitive answer.

The virtual robot's behavior depends largely on all the charts parameters. The smallest change in the chart settings makes major changes in the virtual robot's behavior. The obstacle and workarounds states are particularly sensitive.

A graphical environment and a working programming language in this environment are required for design of such systems. For example, Visual Basic for Application programming language works with the AutoCAD environment.

REFERENCES

- [1] F. Abdessemed, K. Benmahammed, and E. Monacelli. A fuzzy-based reactive controller for a non-holonomic mobile robot.

- Robotics and Autonomous Systems*, 47: 31–46, 2004.
- [2] R. Abiyev, D. Ibrahim, and B. Erin. Navigation of mobile robots in the presence of obstacles. *Advances in Engineering Software*, 41: 1179–1186, 2010.
 - [3] R. Huq, G. K. I. Mann, and R. G. Gosine. Mobile robot navigation using motor schema and fuzzy context dependent behavior modulation. *Applied Soft Computing*, 8: 422–436, 2008.
 - [4] K. H. Lee. First Course on Fuzzy Theory and Applications. Springer-Verlag Berlin Heidelberg, 2005.
 - [5] S. Lee, T. M. Adams, and B. Ryoo. A fuzzy navigation system for mobile construction robots. *Automation in construction*, 6: 97–107, 1997.
 - [6] H. Maaref and C. Barret. Sensor-based fuzzy navigation of an autonomous mobile robot in an indoor environment. *Control Engineering Practice*, 8: 757–768, 2000.
 - [7] E. H. Mamdani and S. Assilian. An Experiment in Linguistic Synthesis with a Fuzzy Logic Controller. *International Journal Man-Machine Studies*, 7: 1–13, 1975.
 - [8] K. Michels, F. Klawonn, R. Kruse, and A. Nürnberger. Fuzzy Control. Fundamentals, Stability and Design of Fuzzy Controllers. Springer-Verlag Berlin Heidelberg, 2006.
 - [9] O. R. E. Motlagh, T. S. Hong, and N. Ismail. Development of a new minimum avoidance system for a behavior-based mobile robot. *Fuzzy Sets and Systems*, 160: 1929–1946, 2009.
 - [10] M. Wang and J. N. K. Liu. Fuzzy logic-based real-time robot navigation in unknown environment with dead ends. *Robotics and Autonomous Systems*, 56: 625–643, 2008.
 - [11] S. X. Yang, H. Li, M. Q.-H. Meng, and P. X. Liu. An Embedded Fuzzy Controller for a Behavior-Based Mobile Robot With Guaranteed Performance. *IEEE Transactions On Fuzzy Systems*, 12(4): 436–446, 2004.
 - [12] L. A. Zadeh. Fuzzy sets. *Information and Control*, 8(3): 338–353, 1965.
 - [13] L. A. Zadeh. A Rationale for Fuzzy Control. *Journal Dynamic Systems, Measurement and Control*, 94: 3–4, 1972.
 - [14] L. A. Zadeh. Outline of a new approach to the analysis of complex systems and decision processes. *IEEE Transactions Systems, Man and Cybernetics*, 1(1): 28–44, 1973.

ABOUT THE AUTHOR

Algirdas Sokas, Doctor, is a Associate Professor at the Vilnius Gediminas Technical University at the Engineering Graphics Department, Lithuania. His research interests are geometry and graphics programming, artificial intelligence. He can be reached by e-mail: algirdas.sokas@vgtu.lt or through postal address: VGTU, Sauletekio al. 11, LT10223, Vilnius, Lithuania.

GEOMETRY AND GRAPHIC DESIGN

Stefano CHIARENZA

Department of Design and Applied Arts - Fine Arts Academy of Naples, Italy

ABSTRACT: The importance of geometry in the analysis and configuration process of the architecture, as well as in the field of industrial or mechanical design, is frequently reminded, recognizing in it an indisputable value in the spatial and figurative reasoning to design and draw architectural or industrial products, even in the age of computers. Less investigated is instead the role of this discipline in the field of graphic design, in which the image is the essence of communication, and the boundaries between image and language assume symbiotic and ambivalent relationships. In this sense, the graphic sign often becomes, itself, an *ekphrasis*, or the narration of contents even as far away from the form. So, in creating an image that must become part of a multi-directional communication, able to drag the eye beyond the usual spaces of experience, it is of the utmost importance the knowledge of the projection's rules that guide, in different ways, all methods of descriptive geometry. In fact, to translate an image into communication it must be designed, organized and controlled, coordinating, in a wise and safe way, the communication and representation tools. And, in this process, the geometry's role is certainly substantial. This paper, on the basis of the teaching and research experience specifically oriented to graphic design, intends to highlight the deep connections between the structuring of the graphic image (lettering, corporate image, brand image, product image etc.) and geometry investigating on significant examples of modern graphic design.

Keywords: graphic design, descriptive geometry, geometrical transformations, image editing.

1. COMMUNICATION BY MEANS OF THE GRAPHIC IMAGE

Communication, intended as a process of imparting or exchanging of information, is one of the key words of several and different fields of study as sociology, politics, economics and involving the culture in general. Human communication is usually given by verbal and written language, but also through non-verbal behaviors or means. Among the nonverbal languages, the graphic communication represents one of the most important ways to create shared understanding, because, by its nature, it creates a common area of communicative commonality. Nonverbal communication, in fact, requires a process of encoding and decoding, in which the generated information can be interpreted from the receiver.

Visual communication is based on vision.

And images, photos, graphic signs and other visual aid constitute the grammar of the so called visual language.

The image has always piqued scientific curiosity of a wide variety of disciplines that studied it as a form of expression also in its relations with the domain of language. However, unlike language, it has never been developed a Science of the Image, but a wide range of philosophical studies have highlighted historical, anthropological and theoreticians aspects that even more underline its conceptual complexity.

In graphic design, the image takes on a central role. Through forms more or less rational it is characterized as cause and effect of a transmission process of ideas, but also as imitation of appearance (*eikon*) or truth (*phantasmata*). The images use their visual resonances to interact with the observer who sees and synthesizes, sometimes in unusual and unexpected

ways, combining sensitivity and intellect, emotion and cognition. In other words, the image shows its content to the spectator's eyes promoting experiences and interpretations [8]. The implications of this, in terms of communication, appear absolutely significant allowing to easily compare the representation to the logos (linguistics proposition): the word in its relation to what it expresses is similar to the way in which figuration is related to the model it represents.

However, if the ontology of the image is like that of the language, it does not have the same structure. While in the language the association between sign and signified is arbitrary, in the visual language there is an intrinsic association: the images are representations that are instead of something else. This involves a substantial limit: the representation by the image becomes an uncertain and changing opinion as the objects to which it refers, even if it has a more intimate and direct link with them.

Therefore the achieving of an image communication depends largely on its abstraction capacity of reality; In other words, on the "...ability to stylize, within the bounded and stable field of the image, the changing perceptual field of the everyday's seeing, with its open borders, its flexible adaptation to different situations" [1]. An anthropological attitude according to some, historical and cultural according to others. However, regardless of the possibility of sharing and reconciliation of the two instances, must be paid attention to the fact that communication depends on the ability to bring out meanings.

So the graphical representation, through a necessary process of abstraction, face each other the magic of the concept and the charm of reality in a fictional representative in which being and appearance, the real and its concept become specular.

In graphic design, the significant expansion of the representing and creative possibilities have given to the image a ubiquitous presence. By means of signs, typography, symbols, figures, drawings, illustration, the use of different

colors, it turns out to be possible to convey a special visual message emphasized, if necessary, with text that can increase signification and meaning.

The use of the image as a means of universal communication and dissemination of knowledge has led to an ongoing research on how the structuring of images that are vehicles of meaning. In this sense, the act of interpretation in the graphic definition of these last is essential both in the definition phase that reading.

To draw a graphic image it is necessary to know the mental mechanisms of perception and reading that govern the response of the same in the mind of the reader.

In the subsequent decoding it comes into play the culture of the single: the reading of the image becomes a subjective interpretation of the viewer, albeit rationalized and guided.

In the complex process of communicative image's structuring is therefore necessary to strictly relate two different forms of logic: that of the designer, based on scientific rigor and objectivity, and that of the observer and his freedom to roam freely with his eyes and thought.

This means being able to combine two ways of understanding the image, two points of view: one who regards his powers of seduction linked precisely to the status of the image, with its implications dreamlike and fantastic suggested to the mind every time the eyes of the viewer investigate on its surface. This type of representation would seem to constitute a privileged space where the observer's desires can be projected, his aspirations and his actual memory and culture. The other one that connotes the image as a rational construction, an intellectual construction governed by the geometry, symmetry and the requirements of perceptual knowledge. In this case, the graphic image is an intelligible model that should be read, interpreted and questioned so as seen [5].

Ultimately, understanding the graphic image as a visual media in the process of non-linguistic communication, it can be traced back to an ar-

ticulated path of conception, design and graphics processing control. The completion of the communicative function with any drafting technique has made it, being the visual transposition of the synthesis of knowledge from diverse disciplines, is inextricably linked to a series of specifically cultural aspects that govern the choices made and that involve multiple fields [2], not only of art and design, but also of the psychology and science of visual communication including that of representation.

2. THE RATIONAL CONSTRUCTION OF THE IMAGE

In the composition of the visual message, the geometry has always been the complementary part of information. Geometry is the understructure of the visual support [7], direct component of the form and the constitutes the process leading to the organization of the communication message and allows to bring the graphic sign to the world of reason by removing the domain of arbitrariness. In the analysis of the visual communication's messages, the opportunity to learn about the rationality of configurative choice appears to be a way to access the creativity. The discovery of proportions, forms, numerical relationships, rational transformations, hidden in the structure of visual figures, partly justifies the design intuition, and reveals, a posteriori, the basic steps in the creative process, even when they appear perhaps not directly sought.

The development of a project is often done resorting to intuitive balance or graphic structures that are not always the result of rules, systems, or preconstituted geometry.

However, understanding and analyzing various aspects of the graphic message, to find some hidden structure that reveals its nature, gives the operator the opportunity for structuring visual images most appropriate to a given communication, working on the aesthetic led of the geometric-mathematical structuring.

The analysis and comparison made on the structure of the graphic image, whether it is in-

tended for the definition of a brand or an advertising campaign, has allowed to highlight the meanings and reasons for forms not always noticeable at first glance, extrapolating, in many cases, connections that address to the intellect first and then to the emotion of the senses.

If the shapes or the basic geometric figures represent the semiotic fundament on which it is articulated the graphic communication, the process of construction of meaning is often elaborated by assembling and modifying the forms on the two-dimensional plane of the drawing.

Among the plane transformations most used are those so-called isometric. They can be seen as part of Euclidean geometry, or as a group of transformations falling in the field of affine geometry, or in wider terms, projective.

The first set of transformations that will be analyzed are those called planes isometries. These transformations give rise to a set of associative phenomena, of growth, of visual rhythm able to assign semantic values to the image that comes with them. In these transformations all the geometric properties of the original figure are invariant under isometry (it preserves measurements, and more specifically distances between points). Among the isometries are classified: translations, rotations, axial symmetries, central symmetries, glide reflexions [9].

Translations. A translation is a transformation of the plane that slides every point of a figure the same distance in the same direction. It preserves orientation, distances, angles. Translation maps parallel lines onto parallel lines and, moreover, a line and its image are also parallel. Translation have no fixed points.

The powers of that transformation in the field of graphic design are significant. Through the translations it is possible, for example, to create patterns and textures sensitizing an area without defining any images but creating a sort of visual attention.

There are numerous designs, in which the translation of basic shapes, combined with a constructed chromatism, create patterns of great

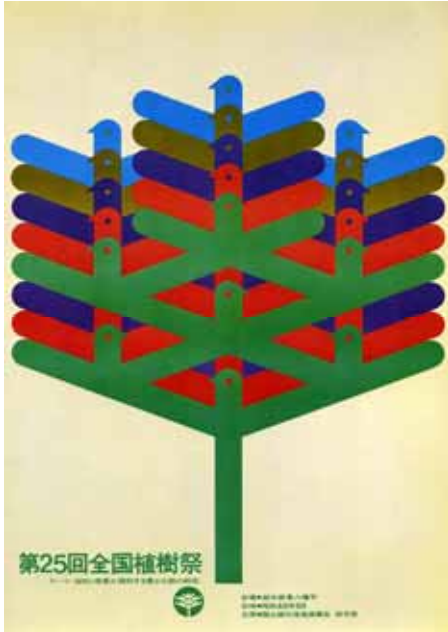


Figure 1: Shigeo Fukuda. Poster artwork for the Tree Planting Festival. 1974

visual impact. Using translations, it may also create the allusion to a figure that will have a different meaning from its original shape. Pregnant examples are two prints: one of the Japanese graphic designer Shigeo Fukuda and the other of the English Alan Fletcher.

In the poster artwork for Tree Planting Festival (1974), the translation of a figure refers to the density of the foliage of a tree. In the drawing of a glass of Beaujolais, however, the simple translation of a circle with alternating chromatism alludes to the wine glass.

In projective geometry, translations can be considered as a particular case of homology with both center and axis at infinity.

Rotations. Fixed in a plane a point O and an angle α , the rotation of center O and angle α is a transformation that to each point P associates a point P' so that $OP \cong OP'$. The angle $\angle POP'$ is congruent to α and has same orientation. So it

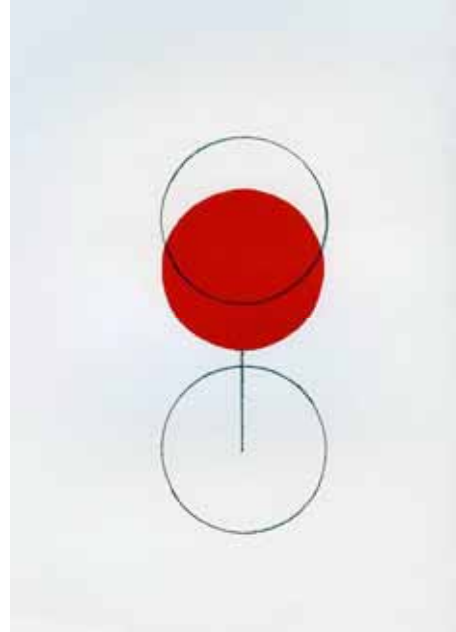


Figure 2: Alan Fletcher. Glass of Beaujolais. Design cover for the review "Idea". 1994.

preserves distances, angle measures, parallelism, collinearity, midpoint.

The rotations, like the translations, are widely used in graphic design. In the application of this mathematical correspondence, often graphic designers exploit the properties of some figures to be self-corresponding respect to particular rotations (i.e. the circle respect to its center, the square to rotations of angles integer multiples of a right angle around its center etc.) associating it with others figures not fixed. In the vast collection of images created for both the definition of logos / brand or for advertising, along with examples, that could be called classics, others appear attractive.

Those of the Italian-Dutch designer Bob Noorda for advertising of the Pirelli Rolle tire, associates to the circular shape of the tire, the rotation around its center of the written in capital letters with different colors. In that way he



Figure 3: Bob Noorda. Pirelli Rolle. Cover. 1959

defines, with the typeface of the Pirelli brand, the wheel, recalling the spirit evoked by the name of the tire. The graphic images realized by Giovanni Pintori in 1963 for the advertising campaign of the Olivetti typewriters possess the same features.

In projective geometry, rotations can be obtained with the product of two harmonic affine homology.

Symmetry. The symmetry can be regarded as an aesthetic-mathematical concept not far from the notions of proportion and harmony that characterized the history of the art's philosophy. As Hermann Weyl wrote "One may ask whether the aesthetic value of symmetry depends on its vital value: did the artist discover the symmetry with which nature according to some inherent law has endowed its creatures, and then copied and perfected what nature presented but in imperfect realizations; or has the aesthetic value of symmetry an independent source? I am inclined to think with Plato that the mathematical idea is the common origin of both: the mathematical laws governing nature are the origin of symmetry in nature, the intuitive realization of the idea in the creative artist's mind



Figure 4: Giovanni Pintori. Advertising print. Olivetti Tekne 3. 1963

its origin in art. I am ready to admit that in the art the fact of bilateral symmetry of the human body in its outward appearance has acted an additional stimulus" [10].

Central symmetry (half turn). Given a point O in the plane, the transformation that associates to a point P another point P' so that O is the midpoint of PP' , is called central symmetry. Point O is the center of symmetry.

Central symmetry has one fixed point (the center), preserves orientation and distance, maps parallel lines onto parallel lines, product of point symmetries in several centers result in a translation, if the number of reflections is even. Two reflections do not commute. The connection between the central symmetry with the rotation and the translation is evident. It is known that the central symmetry is equivalent to a rotation of 180° (half turn). So a double reflection in the same center is an identity transformation.

From this also depends the fact that the figures with a center of symmetry are transformed itself from symmetry around that point (e.g. circle, square, and all parallelograms in general). In graphic design there are many, applications of central symmetries and with very different



Figure 5: Victor Vaserey. Logo for Renault. 1972

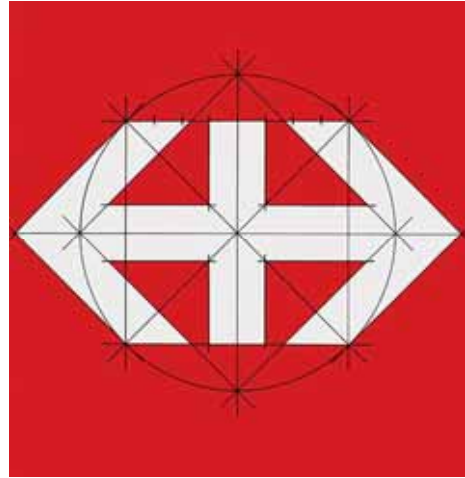


Figure 6: Hans Hartmann. Logo design for SBB Swiss railways, 1972.

results. In projective geometry we obtain central symmetry through an harmonic homothety.

Axial symmetry. Given a line r , the axial symmetry over r is the transformation of the plane in itself that leaves fixed all the point of the line r , and to each point P of the plane associates an image P' so that PP' is perpendicular to r (line of reflection) and cut it in the mid-point.

Axial symmetry is also called line reflexion. Reflection changes the orientation but preserves distances, angles, maps parallel lines onto parallel lines. Successive reflections in two parallel axes result in a translation in the direction perpendicular to the axes to twice the distance between them; successive reflections in two axes that meet in a point O is equivalent to a rotation around O through double the angle between them. Reflexion do not commute.

The geometric-mathematical concept of axial symmetry is one that has the largest aesthetic implications detectable both in nature and in art. And its use in graphic design, operated intuitively or intentionally, gave rise to the structuring of visual messages of great significance and balance. The properties of symmetry have often

been used in fonts or typeface design; sometimes right the symmetry of words has been used for aesthetic purposes.

Glide reflexion. A glide reflexion is an isometry composed of an axial symmetry with a translation parallel to the axis of symmetry. In general a glide reflection is a transformation in the plane that to a point P associates a point P'' finding before P' (reflexed of P around the axis r) then P'' translating P' through a line parallel to that line of reflection. It does not matter whether you glide first and then reflect, or reflect first and then glide.

Many illustrators and graphic designers have been working making evident that mathematical transformation in their works. Escher is certainly one of the most prolific producers of graphic works in which the principle of glide symmetry, collected from the nature's observation and the suggestion of Islamic art, is used for aesthetic and illusory purposes. Are shown below some examples including those of the graphic designer and Charley Harper.

The glide reflexion as well can be obtained as a composed affine transformation (an harmonic orthogonal affinity and an homology



Figure 7: Some artworks of the illustrator Charley Harper.

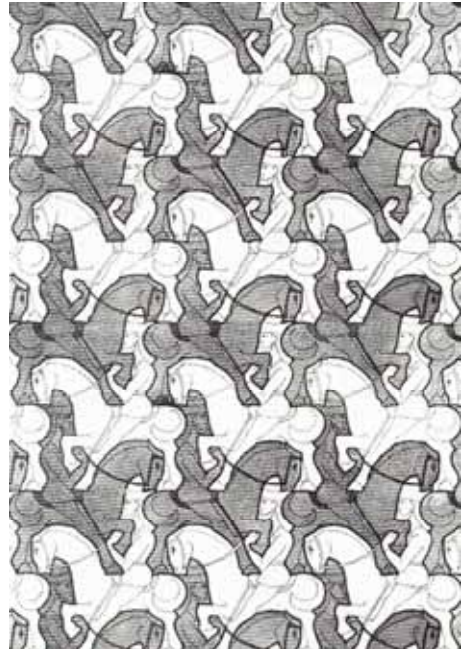


Figure 8: M.C. Escher. Horsemen from a sketches album. 1946.

with both axis and center at infinity in direction orthogonal to the first center).

3. PROJECTIVE TRANSFORMATIONS AND DESCRIPTIVE GEOMETRY AS TOOLS IN GRAPHIC DESIGN

In general, projective transformations on the plane, obtained using the fundamental operations of projection and section, constitute a set of relationships through which it is possible to change figures into others according to a correspondence such as to maintain a set of properties that often allow to recognize the transformed figure even when his metrical characteristics are not preserved. They are called projective properties of figures which include both the graphical and metric-projective properties.

The use of such transformations, using projections and sections, appears to be particularly useful in the field of graphics and this

knowledge is of great utility for designers.

If the geometry in general plays a crucial role in the genesis of the figurative image, it is interesting in particular the graphic power related to the projective transformations of shapes. In addition to the examined cases, which can be also intended as Euclidean isometric transformations (translations, rotations and symmetries) and similarities (translations, rotations and homotheties), the projective transformations allow to produce a collection of signs that have significant mathematical and iconic relationships [4]: in them it is possible to recognize, in fact, despite the obvious and significant variations in shape, an objective expression of the original configuration, conferred by the permanence of the so-called invariant properties.

It is clear, then, that the use of the principles of the projective and descriptive geometry highlights the structural and topological con-

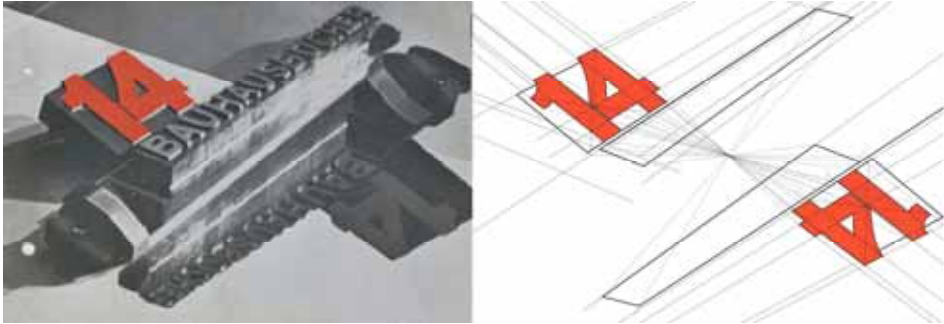


Figure 9: László Moholy-Nagy, 14 Bauhausbücher. Cover 1927.
(Graphics processing by S. Chiarenza)

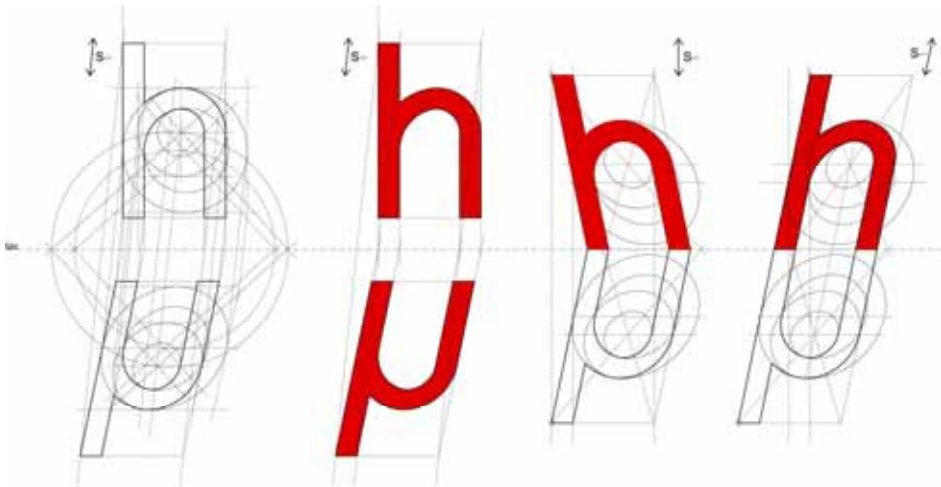


Figure 10: An example of the use of homological transformation in lettering.
(graphics processing by S. Chiarenza)

nections which constitute the very essence of the graphics configurations.

These connections, not related to a metric and quantitative geometry, but to a relational and qualitative one, reveal the complexity of laws and relationships that govern the forms and images and that due to the geometry can be deduced.

So, the search for a structural essence in the visual image created by graphic designer appears as an indispensable approach. The ration-

ality of the image may not be contained only in formal logic. It creeps into its geometric structure that grasps the emblematic aspects of reality, transforming quantitative and sensitive properties into qualitative and relational properties.

As Durand writes the structure can "enhance the notion of 'form' conceived as empirical residue of first instance, and as a semiotic and stiff abstraction, resulting from an inductive process. The form is defined as a kind of arrest, of fidel



Figure 11: Italo Lupi, poster. 1962. The designer shows the use of axonometric projection and shadows theory.

ity, of stillness. The structure implies, on the contrary, a certain transformer dynamism" [3].

In graphic design the role of geometry in structuring of the form can not be then reduced tout court to the science of measure and therefore only to the range of the formal and dimensional properties, expressed by the limited etymological sense of the word. But it should be extended, with its connotations of projective geometry and topology even more, to the ability of formulating more dynamic and generative laws, and for that structural. Through the geometry is not only possible to give substance to the iconic sign, but also the words that often co-exist with it, affecting not so much on the linguistic field as on the structure of the typeface. A logical configuration that avoids the irrational while bringing the mind of the ob-



Figure 12: Herbert Bayer. "Schlesisches Heim" issue 1/2 - 8th year. 1927. Perspective in the cover graphics..

server beyond the limits of the image and language.

Achieving an accurate knowledge of the image's structure and graphic signs that compose it, it means, then, to understand the reasons and internal dynamic laws underlying its development and its possible combinations. A research at the same time analytical and critical that complete the visual experience and offers the designer the possibility to act with full awareness of the external shape of the image in order to seek a deeper objectivity: objectivity that does not stop at the surface, but it is soul and universal substance of the content.

The branding, visual identity, the image of the product are all defined tools in communication with which the graphic designer has to deal constantly, helping to develop and modify the

structuring of these levels of image. Through the image quality actually improves the quality of communication of the underlying assets. Indeed the image, as a means of communication immediately decoded and universally expressive, has the fundamental task of engaging the viewer intellectually. As observed by Kepes "...perceiving a visual image requires the participation of the observer in a process of organization, as the experience of an image is a creative act of integration" [6]. The designer then has the task of controlling the visual language and make a symbolic abstraction transforming the conceptual contents of the communication in a system of signs from immediate and unambiguous decodability. In this research process, the geometrical knowledge plays a key role, allowing to analyze and grasp meanings and translate them into signs, relationships, qualities and functions.

4. CONCLUSIONS

This paper has attempted to highlight the relationship between geometry and graphic design showing, through a set of examples, the importance of using descriptive-geometric configuration processes in the formulation of graphical image for branding, corporate identities, product imaging. Indeed the rational construction of the image, together with the components related to the psychology of perception, is one of the most effective ways to define an image capable of transmitting communication through the strong aesthetic value inherent in its configuration.

REFERENCES

- [1] G. Bohem. Fra l'occhio e la mano. Le immagini come strumenti di conoscenza. In M.G. Di Monte and M. Di Monte. *La svolta iconica*. pp. 212-227. Meltemi, Roma, 2009.
- [2] A.C. Danto. *La trasfigurazione del banale. Una filosofia dell'arte*. Laterza. Roma-Bari, 2008.
- [3] G. Durand. *Le strutture antropologiche dell'immaginario*. Dedalo. Bari, 2009.
- [4] U. Eco. *Trattato di semiotica generale*. Bompiani, Milano, 1998.
- [5] E. Fischbein. The theory of figural concepts, in *Educational Studies in Mathematics*, Springer, 24 (2), 1993, pp. 139-162.
- [6] G. Kepes. *Language of Vision*, Paul Theobald & co. Chicago, 1944.
- [7] B. Munari. *Design e comunicazione visiva*. Laterza, Roma-Bari, 1999.
- [8] H. Shepers. Enthymem. In J. Ritter, *Historisches Wörterbuch der Philosophie*, Schwabe & co, pp. 528-538. Basel, 1972.
- [9] I.M. Yaglom. *Geometrical transformations I*. Mathematical Association of America. New York, 1962.
- [10] H. Weyl. *Symmetry*. Princeton University Press. Princeton (NJ), 1952.

ABOUT THE AUTHOR

1. Stefano Chiarenza is Ph-D. in Drawing and enabled to the functions of associate professor in the field of Drawing. Actually he teaches techniques of space representation for Graphic Design at the Academy of Fine Arts in Naples. His research is aimed at the descriptive geometry in general, with specific interest in the perspective and perspective restitution. He also focuses on the elaboration of digital models and to the representation as a knowledge system of architecture and design.

GEOMETRY AS A TOOL FOR THE DESIGN OF MILITARY ARCHITECTURE: THE EXPERIENCE OF GIOVANNI BATTISTA ANTONELLI

Sandro PARRINELLO¹ and Silvia BERTACCHI²

¹University of Pavia, Italy ²University of Florence, Italy

ABSTRACT: The design of defensive fortresses during the sixteenth century was completed by means of drawing and geometry, by fixing the plan of the pentagonal bastions starting from the line of cannons' enfilading fire and their rifle range. The proportions of the project derived mainly from ballistics studies and geometry was considered to be the essential tool every technician should use to draw up an efficient protection. Giovanni Battista Antonelli, a military engineer of Italian descent and the first member of his important family that worked for the Spanish kingdom for almost a century, explains in his treatise the compositional logics and the necessary proportions used for the design of defensive buildings. His theoretical work on modern fortifications, entitled *Epitomi delle fortificationi moderne di Giovambatta Antonelli*, never published and written in Spain between 1560 and 1561, elucidates the geometric rules to project the elements a modern fortification is composed of. These elements, e.g. ramparts, curtains, moat, embrasures and so on, are described in detail in the first book of the *Epitomi*, specifically dedicated to the defensive architecture, that also gives some information about the different metric units of measurements for the correct dimensioning of architecture, relating the measure in foot units of each country, both in Italy and in Spain. The iconographic setting that enriches the text of the ancient manuscript presents some schematic images that have been examined as well as shaped in a virtual three-dimensional model, in order to provide an exact reconstruction of the "royal bastion" as suggested by Antonelli. The research aims at highlighting the importance of the activities in defensive architecture by Antonelli, and the heritage passed on his family, who disseminated the sixteenth-century architecture in the whole wide world, from Italy to Europe and even to the Caribbean.

Keywords: Spain; 16th century; Giovanni Battista Antonelli; treatise illustrations; geometric analysis; proportional analysis.

1. INTRODUCTION

The sixteenth century military architecture is based on a plurality of disciplines necessary to the engineers to control all the aspects of a defensive project, from the design of the fortress up to the technology for their construction.

However, the topic whose knowledge is considered to be essential, even according to the treatise writers at the time, turns out to be the Euclidean Geometry, that, together with Drawing and integrated by the innovative mathematical erudition of the period, represents

the fundamental design tool for the fortification art and lays the foundations for all the theoretical speculations about the metric value of the fortresses plans [22].

If geometry in civil architecture was used to establish the proportion of orders and ornaments, in the military field it was used to define the plan layout in the fortifications. This was used as a design system that takes into account the variables of shooting firearms, by fixing the plan of the pentagonal bastions starting from the line of cannon's enfilading fire and their rifle range and, at the same time, the require-

ments for proportioning of each element composing the whole system of the fortress.

Through representation there was indeed a control of forms and functions and their links, and the functionality of the defensive buildings was tested by means of experimentation of their geometric characteristics, first of all concerning the plan. The close relationship between form and function was therefore resolvable in a geometric way, with the procedures typical of drawing [11, 18].

2. MILITARY TREATISES AND ESSAYS DURING THE 16TH CENTURY

Starting from the second half of the fifteenth century, especially in Italy where the period of great artistic and cultural splendour of the Renaissance was going on, a significant production of treatises began, in particular relating to architecture and including the military matters. In fact, with the considerable progress of the defensive discipline, it became more and more necessary a specialized tool to summarize the knowledge of defence.

Moreover, the proportions of the project derived mainly from ballistics studies and geometry was considered to be the essential tool every technician should use to draw up an efficient protection.

Therefore there was a period of prosperity for the essays with sections dealing with defensive architecture, focused exclusively on the military topics: the treatises of the art of war. They were written in order to gather the technical arguments and the construction procedures, in addition to operational practices commonly used, mostly translating the concrete experiences of the engineers in their actually built works.

A dual value then, that if on the one hand intended to formulate general properties as a universal reference point, supporting practice with theory, on the other it had inevitably to rely upon pragmatism, adapting the generic rules to the peculiar geographical reality of each situation.

So it is difficult to suppose that treatise writers intended to broadcast the idea of a symbolic and universal model that could come first respect to the constructed work, on the contrary they frequently presented a series of possible cases which fitted to a practically infinite variability of situations and adaptable to real sites, since any designed defensive building first of all had to deal with feasibility and with functionality.

In this context, the treatise had a remarkable value in terms of conveying culture, proving to be a systematic and organized collection of knowledge of a particular discipline.

The main purpose was to convert practical knowledge into theory, that is to say what had been already tested by the writer through experience or new discoveries directly observed.

3. ANTONELLI AND HIS TREATISE ON MODERN FORTIFICATIONS: ALTERNATING FORTUNE OF A COMPENDIUM

Giovanni Battista Antonelli, a military engineer of Italian descent and the first member of his important family that worked for the Spanish kingdom for almost a century, wrote a treatise on modern fortification that fortunately survived to this day, even if never published.

The engineer had been educated in Italy and during his early years he participated to several military campaigns in favour of the Spanish allies, becoming an expert in the field of army quartering and more broadly of defensive architecture. Starting from 1560, he was employed to work for the Spanish Crown as one of the most important king's engineers and technician. He also was a promoter for the other members of his family, who were charged of the defence of the domains, both in Spain and in the overseas colonies, and disseminated the sixteenth-century architecture in the whole wide world, from Italy to Europe and even to the Caribbean [12, 15].

The engineer explained in depth in his treatise, entitled the *Epitomi delle fortificationi moderne*, the compositional logics and the nec-

essary proportions used for the design of defensive buildings, insomuch as the writing represents the main source for the study of the concepts developed by the engineer about defensive buildings and their practical design [21] (Figure 1).



Figure 1: The manuscript by Giovanni Battista Antonelli, presently preserved in Toledo, Spain (one of the pages of the “*Epitomes on artillery and firearms*” in Book I, folio 54).

Composed of three well-defined and independent works, all written in Italian, the treatise deals with some of the most significant aspects of the military field: fortification, artillery and army quartering.

The work was written up in Toledo after the Treaty of Cateau-Cambr  sis that ratified in 1559 the peace between Spain and France after their longstanding belligerence [20].

The essay is prior to any defence project designed or built by the engineer, hence proposing only Antonelli’s cultural baggage at the time of his relocation to Spain, showing his cultural education in the military science of Italian origin.

In the work the engineer gathers together the experience of his recent participation to the sieges and belligerences carried out during the Spanish military campaigns in Italy and Flanders, recording in the chapters a fair amount of information with the state of the art of the discipline.

The conciseness of the text, as stated by the title, was due to the fact that this was the shorter edition of a previous wider essay, got lost some years before, and moreover it had to be subject to an additional editing at a later time.

The merely notional nature of the work and the teaching approach of the author make the manuscript more as a manual of practical use than as an innovative theoretical essay for the discipline. Nevertheless, the engineer demonstrates to be a good expert both in the theoretical and practical modern defensive techniques, revised on the basis of his personal careful consideration and in the light of the new geographical setting where he advanced his professional career [3].

In fact, in the treatise there is a preview of the design methods and defensive strategies that would have been adopted to the Spanish reality and of all the innovative concepts of military defence for the kingdom, empirically verified by the engineer in later times.

On the contrary the “ideal model” for the architectural design of the fortification is not revealed within the text, and in the real circumstances it is always adapted to the site characteristics, however the compositional models based on the proportions and the drawing rules are provided indeed [16].

It stands to reason that the book has a promotional nature, which main objective is to show Antonelli’s grounded technical skills in the presence of the upper rank of military and political administration, in hopes of entering into the monarch’s good graces and thereby to ensure his personal professional career.

So it is no coincidence that, as custom at the time, the three writings were dedicated to the King Philip II and to the general captain of the imperial artillery, Juan Manrique de Lara, with whom the engineer had started collaboration since his first years attending upon the Spanish Crown.

The cultural level of the books is profound, with recurring erudite quotations from classical authors especially Vitruvius (1st century BC)

and Vegetius's *Epitoma rei militaris* (5th century AD) [7].

The setting of the writing is similar to the renowned treatises at that time, and although there are no direct references some affinities in the texts may suggest a probable contact with other important coeval authors, such as Pietro Cataneo [9] and Giovanni Battista Zanchi [23].

Regardless of the criticisms on the complete originality of the concepts contained in Antonelli's treatise, it is important to highlight that the essay had a crucial importance as a mean of cultural diffusion between Italy and Spain.

In fact, despite it had never been published and the limited diffusion among the coeval engineers, the fundamental contribution was spread worldwide through the works and projects by the members of his family [10].

3.1 Structure and distinguishing features of the *Epitomi*

The treatise by Antonelli was part of the collection of the Museo del Ejército in Madrid since 1960, but it is presently preserved at the Instituto de Historia y Cultura Militar in Toledo.

It is a well-preserved manuscript consisting of 147 sheets, written up with brown ink on both sides of the page and without marginal notes, bound with a parchment cover. The text, with legible calligraphy, is a copy of the three different works the *Epitomi* are composed of.

The language used by the author is correct, except for some dialect or sporadic use of Spanish terms, and the sentences mostly follow a non-linear narrative form. The drafting of the text fully reflects the Renaissance culture dealing with the matters of war.

The dating of the writings is deducible by some historical fact referred in the text that delimits the writing in the period from the 24th of April 1560, up to March of the following year.

In the volume the three works appear without regard to the chronological order in which they had been written: first of the *Epitomi delle fortificationi moderne di Giovanbatta Antonelli* (ff. 2-41 r), then the *Epitomi del Trattato*

dell'Artiglieria (ff. 41 v-122 v) and finally the *Epitomi della maniera di alloggiare un campo* (ff. 123 r-146 v).

From the comparison of the information contained in the texts we can assume that the section about the army is the first work by Antonelli, completed in the spring of 1560; hereafter he wrote the essay on modern fortifications and afterwards, during the first months of 1561, he started the substantial section relating to firearms.

The writings, each built on a specific topic of the military subject, have a repeated structure, though maintaining an independent nature, and follow the traditional organization of the treatises: after the dedication and the prologue, the general table of contents precedes the extended discussion of the various chapters.

The iconography setting presented in the text is generally of fair quality. The images, drawn with brown ink, complete the text for better understanding and with illustrative purposes. The coloured backgrounds have different meanings: as a simplification that links to each colour a specific function, or for the indication of the sectioned walls in the case of the fortification section. They have, depending on the section, a very different level of detail, which can present very simple sketches or the meticulous representation of objects embellished with details and nuances of colour.

3.2 Sections of the treatise

In the first section of the treatise, the *Epitomi delle fortificationi moderne*, Antonelli elucidates the basic precepts for the defence of a city according to the concepts of modern fortification. This essay, written in 1560, is dedicated to the eminent Juan Manrique de Lara, the then-general captain of the Spanish artillery.

The writing is organized in more than forty paragraphs, dealing with the theme of the defence of cities and fortresses. After an introduction on the general concepts for the choice of the best site where locating a fortified city, all the architectural elements that compose a for-

tress are analysed in detail, providing dimensional information useful for the project. The last part of the text focuses on the different types of fortified settlements, presenting exhaustively different case studies.

Moreover, Antonelli writes about the technique of rammed earth both from the point of view of materials, besides technological and constructive process. The engineer suggests its use to build walls and bulwarks, because of the possibility for earth to contrast efficaciously firearms, resisting better than any other masonry to the impact of cannonballs.

In this first section of the work the text is supplied with twelve illustrations, all drawn with brown ink and filled with dark brown shades in order to identify the sectioned elements. The images are not always proportionate and often without any metric information, referring for measures to the tables in the text.

The figures are essential, didactic and lacking in detail, except for a drawing of the plan of the bastion, very similar to the later project by Antonelli's most famous brother Battista for the defence of the city of Peñíscola, Spain.

The author prefers the plan representation for the illustrations and uses neither the profile or elevation view nor the three-dimensional representation, apart from the only case of the bastion illustrated in figure 2, where the axonometric representation is used to show a correct and complete view of the object.



Figure 2: Axonometric representation of a bastion, *Epitomi*, folio 22 v.

Dealing with the topic of firearms, the *Epitomi del trattato dell'Artiglieria* is the larg-

est section of the treatise, counting even 81 sheets. Last work drafted by Antonelli on the basis of a chronological order, the writing starts from the back side of the sheet number 41.

According to Antonelli's information, the draft had been completed in just two months and ended on the 11th March, 1561. The final version ought to be subject to further additions and reviews that never took place.

The dedication to King Philip II demonstrates that the engineer wanted to show the sovereign his in-depth examination of one of the most important and relatively innovative subjects at the time. The knowledge of artillery, in fact, was essential for technicians since fortifications were exactly calculated to resist to the destructive force of firearms.

The introduction to the section starts with four refined sonnets, highlighting the engineer's military skills. The table of contents of the three books show several paragraphs, some of them never written, but all the topics are not organized on a same theme, being the whole section very fragmentary and chaotic.

The discussion concludes with two simulated attacks and battles that are won by the General Captain of the Artillery, key figure of the whole dissertation.

This section counts 22 full-page illustrations, in the majority concerning the first and second book. Some of the images refer to missing paragraphs respect to the initial table of contents. The pictures, that have their own numbering, were subjected to a new ordering subsequent to the trimming, and not always respecting the logical dissertation of the text, i.e. figures referring to the first book are included in the second one and vice versa.

The illustrations depict several objects, coloured with decorative intent and characterizing the materials they are made of, even though always lacking of references, such as letters or a legend.

The third section of the treatise dedicated to the army quartering, entitled *Epitomi della maniera di alloggiare un campo*, was written in

Toledo and concluded on April 24th, 1560.

Dedicated to Captain Manrique de Lara, this is the shortest writing of the volume, composed of only 23 pages, and the first to be drawn up in chronological order.

In the prologue Antonelli states the reasons why he wrote about his military experience, working on it in a period of political calm when he was free from military service. He wanted to prove his knowledge and practice as a soldier after the battle of Dorlano (1557), when he attended upon the General Master Jean de Ligne. The event ended quickly, preventing the engineer from demonstrating his competence.

In this last section of the treatise there is no table of contents but only one single extended chapter, organized in different topics.

With regard to the key themes, the safety of an army quartering depends on its location, defended by natural features or by artillery and good practice, so in this case the General Captain and the engineer have the major responsibility.

The most interesting contribution of this part of the treatise is the enumeration of ten case studies of camp settlements, determined by the geographical variables of the site. They are represented by as many very interesting illustrations about various military camp settlements in different sites. The drawings, full-page and on the only front side of the sheet, are made in watercolour, with yellow, red, ochre, blue and green shades. They are attributed to Antonelli, who only after starts his collaboration with the painter Antón Coll [8].

The representations have a fairly accurate delineation of the military camps, showing at the same time both the fortifications and the boundaries in the plan, while the profile is used to draw the encampments. The reference letters, existing only in the first five illustrations, are explained in a legend and also the colour provides additional information about functions.

In general, the drawings are lacking of metric scale, although units and numerical notation are written in the text ("pace of walking man"),

and the North is not explicitly shown.

With regard to the representation in general, there is a pleasant aesthetic connotation of the surroundings landscape in the figures. Graphic details representing natural and artificial elements (such as rivers, trees, forests, flowers, hills or human elements, like roads, settlements, residential buildings, churches and chapels, both wind- and water- mills) complete the drawing of the typical elements of the military outposts, such as towers, walls and ramparts.

4. DESIGN RULES BY ANTONELLI

The elements typical of modern fortification, e.g. ramparts, curtains, moat, embrasures and so on, are described in detail in the first book of the *Epitomi*, specifically dedicated to defensive architecture, that also gives some information about the different metric units of measurements for the proper dimensioning of architecture, relating the measure in foot units of each country, both in Italy and in Spain.

First of all, Antonelli describes the general precepts for designing a defensive building by means of geometry.

A proper design starts from the choice of the plan shape of the fortress, better when closer to the circular path, and of its parts, in proportion- al connection among them (Figure 3).

Scholars in fact consider inappropriate the use of triangular and quadrilateral shapes for the plan because a few sides polygon determines acute angle bulwarks for the fort, considered less resistant to firearms (Figure 4).

The reasoning is above all based on geometrical verification, since the angular size of the bulwark is bigger when the plan has a large number of sides, and as a consequence as the inner angle between two curtains increases and becomes obtuse [13] (Figures 5-6).

Nevertheless, the more complex plan shapes are infrequently applicable to real buildings, first of all because they are more expensive and also due to the peculiar geographical features of each location.

In Antonelli's treatise, differently from what

happens in the coeval essays on military architecture, there is not an initial section completely dedicated to geometry, perhaps due to the concision of the work, or maybe because the knowledge of the subject is assumed among the experts. It is obvious that the engineer uses the rules of geometry as tools to create a project, but he does not intend to repeat these demonstrations in a specific way within the book.

There are only two mentions to geometric constructions in the first *Epitome*, dealing with the design of a generic bulwark.

The project of the plan is none other than the operational application on the paper of right angles, parallel and endless lines, and intersections in relation to the vision lines and shooting trajectories, starting from the opposite embrasures along the curtains and walls (Figure 7).

Antonelli suggests, with the aim to draw the plan of an architectural element, to measure a distance of 120 or 130 feet starting from the inner virtual corner of the curtain, and to consider a perpendicular line that represents the flank and the protrusion of the side.

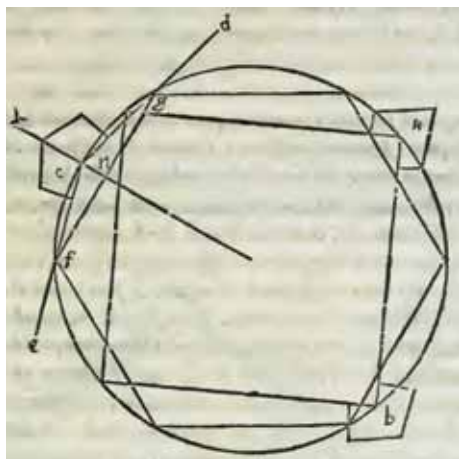


Figure 3: Geometrical demonstration by Lanteri about the widening of the angle of the bulwark respect to the increase of the sides of the fortress, page 28 [13].



Figure 4: Schemes of quadrilateral, pentagonal and 14-sided fortresses, inferable from the text by Antonelli.

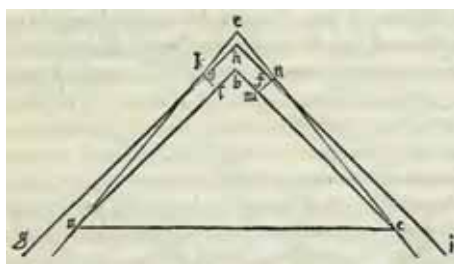


Figure 5: Demonstration by Lanteri that given a right angle between the curtains, it is possible to draw only a bulwark with an acute angle, page 25 [14].



Figure 6: Demonstration by Tartaglia about the bastion functionality when altering his angle respect to the two connected curtains (smaller, equal or bigger) [19].



Figure 7: The sketch of a bastion by Antonelli with the moat determined by drawing endless lines based on the artillery shooting. Image at folio 16 r.

Considering as a bench mark the opposite bastion and passing through the end point of the side wall, it is possible to define the extremity of the bulwark, by intersecting the two symmetrical endless straight lines.

In the same way it is recommended to define the boundaries of the moat, that is to say by means of an endless line starting from the inner corner of the flank and intersecting in front of the bulwark extremity and in the middle of the curtain, creating a zigzag path. Therefore, once chosen the first dimension, any element of the system turns out to be defined by applying the rules of proportioning, with the same method also for sections and elevations.

The iconographic setting that enriches the text of the old manuscript presents also some schematic images that have been examined as well as modelled into a virtual 3D environment, in order to provide an exact reconstruction of the “royal bastion” as suggested by Antonelli (Figures 8-9-10).



Figure 8: Original drawing, folio 16v.

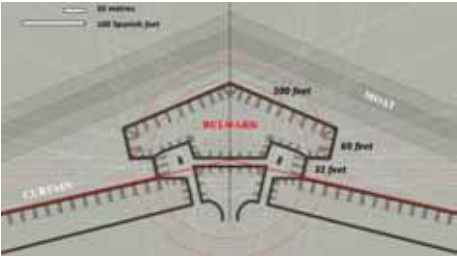


Figure 9: Proportional scheme as described in the first book of the treatise. Drawings by S. Bertacchi

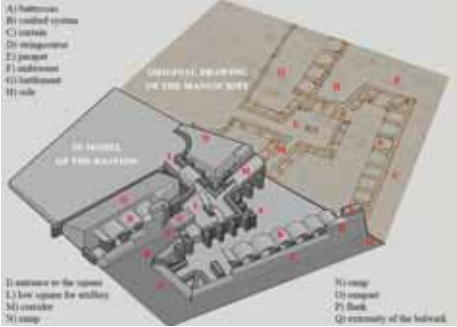


Figure 10: 3D reconstruction of the bastion in comparison to the original drawing of the manuscript. Original document at folio 16 v of the treatise. 3D model by S. Bertacchi.

5. MODEL OF THE TOWER AND GEOMETRY AS DESCRIPTIVE SYSTEM

One of the most important geometrical models used in this period for the design of towers is the hexagonal one, which represents the extreme result of the application of regular geometric shapes to reality, anticipated in the treatise for what concerns the building material [17].

In the Kingdom of Valencia there were only two towers of this kind, now completely destroyed but recorded in a report dating back to 1580 [6].

According to the archive documentation, this model had been conceived during the reconnaissance mission of the Kingdom of Murcia in 1570, made by Vespasiano Gonzaga together with Giovanni Battista Antonelli.

The nobleman thought that the hexagonal shape allowed a better management of firearms and offered more space for the movement of the cannons compared to the square-shaped and even the circular one. In fact the hexagon was the middle way between the resistance of the circle and a polygon with a sufficient number of straight sides, compatible with the expansion of the upper parade ground.

Probably Antonelli actively contributed to the formulation of the model, which could be

the result of the exchange of opinions on the subject of the two engineers on the best shape for defence.

The same model was adopted by Antonelli in 1578 for the construction of the towers in the area of Lorca (Murcia, Spain).

There are two important graphic documents survived to this day and attributed to the engineer: the first one provides a synthetic scheme with proportions for the construction of a six-sided tower in rammed earth (Figure 11; compare with the digital model in Figure 12).

The second one represents the tower's plan and cross-section in an orthogonal projection with in evidence the thickness of the walls with the scale in Spanish feet, the *pies* (Figure 13).

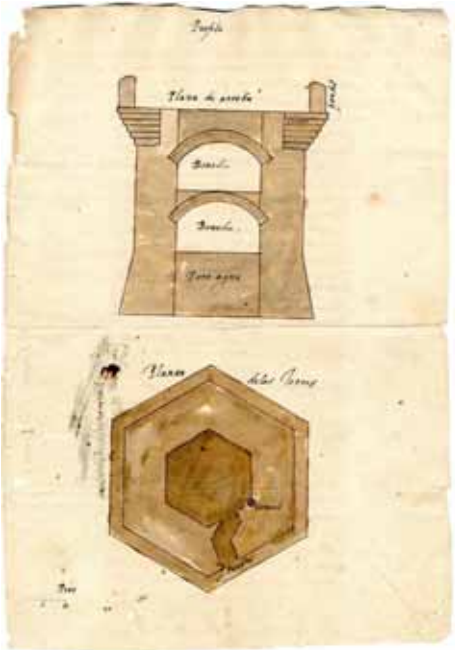


Figure 11: Plan and elevation of a hexagonal tower built with rammed earth technique. Document attributed to G. B. Antonelli. *Archivo Municipal de Lorca, Spain, Legajo monográfico torres de la marina, year 1578.*

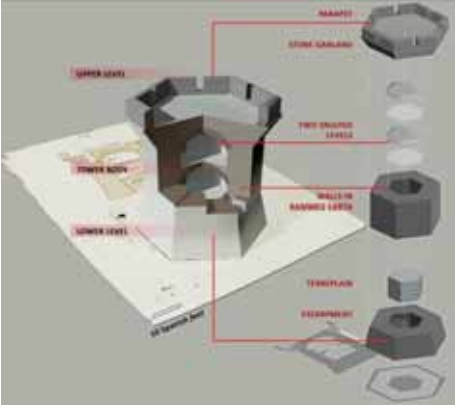


Figure 12: 3D modelling of a hexagonal tower and its elements from the drawing in orthogonal projections. Model by S. Bertacchi.

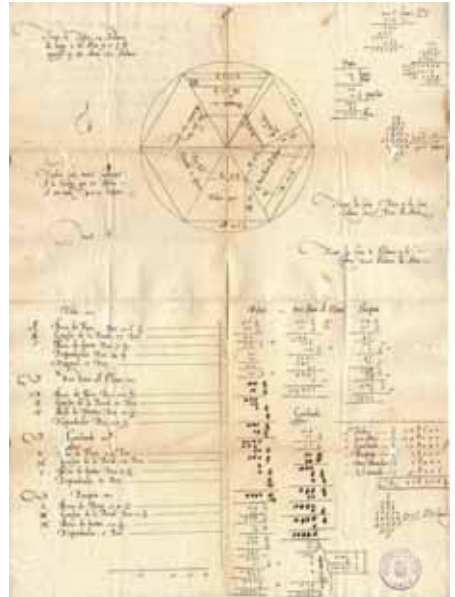


Figure 13: Graphic model for the proportions of the elements of the hexagonal tower. Document attributed to G. B. Antonelli. *Archivo Municipal de Lorca, Spain, Legajo monográfico torres de la marina, year 1578.*

In the general schematic plan, within each side of the hexagon there is all the information necessary to the development of the project: drawing is used to determine any proportion, on a specific scale, ready for the generic application to each element of the front elevation.

In the first slice of the polygon the radius of the base of the tower is equal to 26 feet and a half (A) and the wall thickness for a total amount of 13 feet (B). The inner empty space is 11 feet and a half (C-F); the lateral sides of the hexagon are 10 feet in cross-section (E). The garland sticks out 3 feet (H) and the parapet (thickness 2 and a half feet) encloses an upper square for a total of 43 feet (Figure 14).

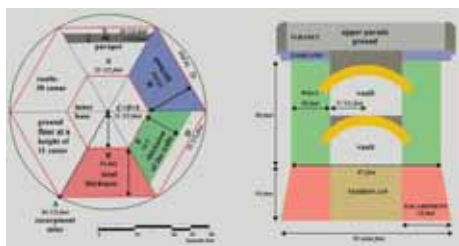


Figure 14: Study of the proportions of the building plan and relative cross-section according to the project documentation.

Each “tapia” made of rammed earth has the dimensions of 14 x 5 x 2 and a half palms and a total volume of 175 cube palms, useful to calculate the necessary materials for the construction work.

In the report related to the drawing the elevation dimensions and the construction technique are profusely specified in the text.

The slope of the basement has to be one to five, in this way decreasing the size of the wall of 3 feet for a total height of 15 feet, level of the access.

An inner spiral staircase allows going up to the upper floors, barrel vaulted with a brick structure. In the project the walls have to raise the height of 28 feet up to the garland (Figure 15).

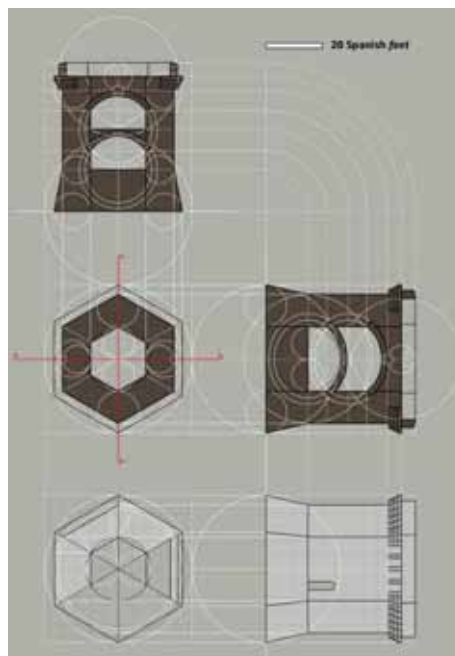


Figure 15: Two-dimensional reconstruction of the tower starting from the only indications of the generic design by Antonelli and research of the proportions among the constitutive elements. Plan, elevation and cross-sections, drawing by S. Bertacchi.

6. CONCLUSIONS

The experience of Giovanni Battista Antonelli in the field of military architecture was grounded on the common knowledge spread by many engineers at the time, but with a personal interpretation. His theoretical contribution, the treatise on modern fortifications, once again proves the interest of the period in using geometry as a tool to design defensive buildings and to check the compositional logics for the proportioning of the fortifications [4].

Unfortunately the lacking publication of the *Epitomi* determined the impossibility of a wide dissemination of the treatise, and the planning practice by Antonelli was spread mainly

through the written reports or the projects made by the engineer, rather than through his essay. Except for the family of Giovanni Battista, who probably examined the manuscript or at least exchanged advices with each other on the topic, the popularization of the treatise might have been limited to very few readers.

Probably the copy of the manuscript survived to this day, or maybe one of the previous drafts was consulted by Captain Manrique de Lara who, according to the author, revised the contents before the presentation to King Philip II.

One cannot exclude that the sovereign himself consulted at least one section, as it might be thought thanks to a hint in the prologue, which refers to his personal knowledge of the work.

Among the following writers and scholars, it is very likely that the engineer Cristóbal de Rojas had known the volume of the *Epitomi*. In fact, the Spanish captain at the end of the 16th century deals with some issues in a similar way respect to the work of Antonelli, presumptively consulting one of the manuscript copies in circulation. For example, both authors agree with the defence net proposed for the protection of the entire Kingdom of Spain [5]. In the *Epitomi delle fortificationi moderne*, Antonelli compares the fortification of the whole kingdom to the one relating in a smaller scale to the city defence, assimilating the coastal military outposts to the ramparts of the city [1].

The same concept was recommended after nearly forty years by De Rojas, with just alike words, that one more time proves the innovation of the military ideas by Antonelli.

In conclusion, one can say that, although the manuscript was never published, Giovanni Battista's models were disseminated by his family members all around the world, carrying on with his military architecture [2].

ACKNOWLEDGMENTS

Acknowledgments to Mr. Manuel Muñoz Clares and to the Archivo Municipal de Lorca, Spain, for the graphic documents on Murcian towers attributed to G. B. Antonelli.

REFERENCES

- [1] G. B. Antonelli. *Discurso sobre la fortificación y defensa del Reyno de Valencia del maestre racional de aquel Reyno, y de Juan Bautista Antoneli*. Archivo General de Simancas (Spain), Estado, I, f. 13, 1563.
- [2] S. Bertacchi. *Modelli compositivi per la difesa "alla moderna". L'esperienza di Giovanni Battista Antonelli* (PhD thesis). Dipartimento di Architettura, Firenze, 2013.
- [3] S. Bertacchi. L'arte fortificatoria dell'ingegnere italiano Giovanni Battista Antonelli: un'analisi fra teoria e pratica militare tardo cinquecentesca. In L. Carlevaris (edited by). *Linee di ricerca nell'area del Disegno. Approfondimenti dalle tesi di dottorato in Mostra*, pages 113-122. Aracne Editrice S.R.L., Roma, 2013.
- [4] S. Bertacchi. Compositional models by Giovanni Battista Antonelli and his experience in the military field. In *Nexus 2014 Ph.D. Day presentations, June 9-12, Ankara TR*, pages 80-87. METU, Ankara, 2014.
- [5] J. V. Boira Maiques. Geografia i control del territori. El coneixement i la defensa del litoral valencià al segle XVI: l'enginyer Joan Baptista Antonelli. In *Cuadernos de Geografia*, 52: 183-199, 1992.
- [6] J. V. Boira Maiques. *Las torres del litoral valenciano*. Conselleria de Infraestructuras y Transporte, Valencia, 2007.
- [7] A. Cámara Muñoz. Las torres del litoral en el reinado de Felipe II: una arquitectura para la defensa del territorio-I. In *Espacio, Tiempo y Forma: Revista de la Facultad de Geografía e Historia*, 3, (VII): 55-86, 1990.
- [8] A. Cámara Muñoz. La profesión de ingeniero: los ingenieros del rey. In *Técnica e Ingeniería en España I, el Renacimiento*, pages 125-164. Ed. Manuel Silva Suárez, Zaragoza, 2004.
- [9] P. Cataneo. *I primi quattro libri dell'architettura di Pietro Cataneo senese* (...). Venezia, 1554.

- [10] F. Cobos Guerra. Ingenieros, tratados y proyectos de fortificación: el trasvase de experiencias entre Europa y América. In P. Chías and T. Abad (edited by). *El Patrimonio fortificado. Cádiz y el Caribe: una relación transatlántica*, pages 173-210. Universidad de Alcalá, Madrid, 2011.
- [11] J. Galindo Díaz. *El legado constructivo de los tratados de fortificación. Siglos XVI-XVIII*. Impreso Universitario Serie Cuadernos de Investigación, Universidad del Valle, 6, 2000.
- [12] G. Gasparini. *Los Antonelli. Arquitectos militares italianos al servicio de la Corona española en España, África y América (1559-1649)*. Venezuela, 2007.
- [13] G. Lanteri. *Due dialoghi (...)*. Bolognino Zaltieri, Venezia, 1559.
- [14] G. Lanteri. *Duo libri (...)*. Bolognino Zaltieri, Venezia, 1559.
- [15] A. Llaguno y Amirola and J. A. Céan Bermúdez. *Noticias de los arquitectos y arquitectura de España desde su restauración*. Madrid, 1829.
- [16] S. Parrinello and S. Bertacchi. Giovanni Battista Antonelli: theory, model and reality. In S. Parrinello, S. Bertocci and G. Pancani (edited by). *Between East and West: Transposition of cultural systems and military technology of fortified landscapes, Congress Proceedings, Poppi (AR), May 7-13, 2012*, pages 184-187. Edifir Edizioni, Firenze, 2012.
- [17] S. Parrinello and S. Bertacchi. Il disegno delle architetture militari. Le torri costiere spagnole del Cinquecento. In *Actas XI congreso internacional de expresión gráfica aplicada a la edificación*, page 1079. Universitat Politècnica de València, Valencia, 2012.
- [18] G. Severini. *Progetto e disegno nei trattati di architettura militare del '500*. Pacini editore, Pisa, 1994.
- [19] N. Tartaglia. *Quesiti et inventioni diverse de Nicolò Tartaglia (...)*. Venezia, 1554.
- [20] M. Sartor (edited by). *Giovan Battista Antonelli, Epitomi delle fortificationi moderne*. Forum, Udine, 2004.
- [21] M. Sartor (edited by). *Omaggio agli Antonelli. Considerazioni intorno a tre generazioni di architetti militari italiani attivi nel Mediterraneo e in America, Atti del Convegno Internazionale di Studi (Gatteo 3-5 ottobre 2003)*. Forum, Udine, 2004.
- [22] A. Vera Botí. *La arquitectura militar del Renacimiento a través de los tratadistas de los siglos XV-XVI*. Universidad Politécnica de Valencia, Valencia, 2001.
- [23] G. B. Zanchi. *Del modo di fortificar le città (...)*. Plinio Pietrasanta, Venezia, 1554.

ABOUT THE AUTHORS

1. Sandro Parrinello is architect and researcher at the Department of Civil Engineering and Architecture at the University of Pavia (Italy) and European Ph.D. in *Science of Surveying and Representation of Architecture and Environment* at the Department of Architecture of the University of Florence (Italy). His research aims at representing landscape by means of Drawing. Since 2005 he has been carrying out the research and the survey of the fortresses designed by the Antonellis all around the world, especially in the Caribbean area.
2. Silvia Bertacchi is architect and European Ph.D. in *Science of Surveying and Representation of Architecture and Environment* at the Department of Architecture of the University of Florence (Italy). Her research is mainly focused on Descriptive Geometry and its application on architecture. Since 2010 for her PhD thesis she has been studying the compositional models in the field of 16th c. military architecture, in particular designed by Giovanni Battista Antonelli, deeply analysing the geometrical tool used for defensive buildings.

THE GEOMETRY BEHIND THE OCTAGONAL HALL OF SMALL THERMAL BATHS AT HADRIAN'S VILLA

Luca CIPRIANI¹ and Filippo FANTINI²

¹Alma Mater Studiorum-Bologna, Italy ²Alma Mater Studiorum-Bologna, Italy

ABSTRACT: Hadrian's Villa was a place of architectural experimentation, wanted by the Emperor Hadrian during the 2nd century AD in order to express a very personal and revolutionary idea of art, of architectural design and in particular his new formal conceptions of vaulted spaces. The first ribbed vaults of the Roman Empire were conceived there, at his Tiburtine villa, where the incessant sequence of promenades and naturalistic environments culminated into innovative vaulted spaces. Some of those pavilions, at that time, must have appeared so strange and futuristic that Apollodorus of Damascus, famous architect of the previous Emperor Trajan, did not appreciate their shapes, in his opinion more similar to "pumpkins" rather than domes (Cassio Dio, Roman History, 69, 4). Among the numerous domes of the Villa, one in particular, still adequately preserved, represents a unique example in the history of architecture, since its rare tent-shape no longer appeared in other Roman and Late Antiquity examples. This dome is the one that covered the octagonal hall of small thermal baths of the Villa. Thanks to new surveys, carried out through laser scanner and processed by means of reverse modelling applications it was possible to obtain highly reliable contour lines and sections able to enlighten some geometrical aspects of this dome that still continued to be misunderstood. An original and practical geometric construction was aimed at linking together the drawing of plan (*ichnographia*) and section (*orthographia*) of this octagonal mixtilinear hall, showing a not common skill of the architect in taking advantage of geometrical properties. The result was a vault formed by a thin, continuous surface, as the one of a tent. Starting from a series of relevant sections obtained by the high resolution mesh model of the vault, then filtered and idealized after to a in depth metrological analysis, it was possible to achieve a new solution for this original artefact. In order to generate a continuous surface starting from contour lines of a partially preserved and highly deteriorated vault, whose central part fell down centuries ago, it was necessary to use subdivision surfaces rather than NURBS for the integration of missing parts. It seems probable that this daring solution was applied to other two buildings inside the Villa: the southern hall of the Piazza D'Oro and the access to the Accademia. Both these spaces present a common initial octagonal shape, modified in order to obtain a variety of striking halls. At Hadrian's Villa there are many domes, some of them are characterized by more traditional shapes (semi-spherical and basins), others are ribbed vaults, and some others are intermediate solutions. But the singular "tent" of the Small Thermal Bath, perhaps the most original, had no luck and did not become a formal archetype for the later periods, in contrast to what had happened to the domes of the vestibule of the Piazza d'Oro and of the Serapeum.

Keywords: Villa Adriana, *Ichnographia*, *Orthographia*, Subdivision Surface, Reverse modelling.

1. FROM VILLA'S GENERAL DESIGN TO SINGLE PAVILIONS

Hadrian's Villa was built from 118 until 134

AD with the aim of being both an extensive mansion and seat of government [2]. Many authors focused their interpretation on the nature of this Villa as a small city, a sort of Empire in

miniature [10]. Actually, the scale of Hadrian's mansion is even bigger than a town like Pompeii; the main difference, related to the topic of this paper, was that the Emperor wanted to create a sort of ancient "World's fair", where skilled architects from the whole Mediterranean Basin had to develop a vast environment of pavilions. Hadrian wanted to exhibit in his grand palace a series of buildings, characterized by new shapes and constructive techniques: real prototypes for centuries to come.

Many authors belonging to the field of architecture and archaeology [11, 6, 3] suppose that the majority of the Villa was entirely planned at the beginning of the construction, with some minor work-in-progress modifications.

In this paper another theory is embraced, deeply related to the interest of the Emperor with regards to vaulted spaces and in particular central plan pavilions, whose distribution into the general layout of the mansion represents a fundamental reading key for the whole complex and its development.

Hadrian's personal interest in ribbed vaults is widely known, since the anecdote quoted by Cassio Dio about his quarrel with Apollodorus of Damascus, Trajan's favourite architect, is centred on some drawings (or models) of peculiarly-shaped domes. Apollodorus compared their design to pumpkins [4]: "Be off, and draw your gourds. You don't understand any of these matters." The Syrian architect considered Hadrian's proposals no more than amateurish, odd "experiments", but on his accession, finally he could put into practice his new aesthetic in the frame of the Tiburtine dwelling.

Being a place for both institutional tasks and convivial activities the villa was the perfect scenario for the creation of astonishing novel constructive solutions, mainly expressed by two uses of vaulted spaces:

- Semi-domes as convergence points at the end of long promenades (for instance the so-called Serapeum).
- Astoundingly discoveries of huge hidden

domes after a sequence of interiors and courts (southern hall of the Piazza d'Oro, Reverse-Curve pavilion of Accademia).

The hypothesis suggested here deals with two possible aspects in which vaulted spaces of Hadrian's Villa may have played a fundamental role as "milestones" upon which the architects developed the rest of the buildings that forms the characteristic aspect of Hadrian's Villa. This last assumption is particularly true in the case of pavilions aimed at public and institutional functions, including recreational and entertainment aspects of the mansion's life.

The first aspect to point out is the presence of buildings characterized by similar functions and shapes that, starting from the first nucleus of the Villa and built approximately between 118 and 121 AD [2], are duplicated toward the two main expansion lines on the flanks of the hill during later phases (125-128 AD and 133-134 AD). This tendency seems to underline the intentionality of the Emperor toward a constant renewal, enrichment and enlargement of *triclinia*, thermal baths, nymphaeums, etc. [5].

The second aspect concerns the necessity of the Villa's ancient architects to carry out an experimental phase for testing Hadrian's design hints and wills. Salza Prina Ricotti [13] identified inside the complex of Great Thermal Baths a possible area aimed at design studio: there it has been discovered a marble model of a building for spectacle, that witnesses with other finds the presence of a constant design activity.

Hadrian gathered inside his Villa a huge court of different professionals, ranging from sculptors, architects, musicians, mathematicians and geometers; they came from different areas of the Empire and some of them were designated to solve difficult constructive and designing problems.

During nearly twenty years, necessary to the construction of the Villa, many different solutions were tested, stressing the new structural possibilities achieved by the technology of massive concrete vaults developed during the thirty former years by Roman engineers [10].

Inside the Villa coexist domes and half-domes of different shapes ranging from more traditional ones (hemispheric) to ribbed vaults and tent-like ones. Some of them present a completely new design and differ in comparison to other contemporary structures (like the Pantheon) for being extremely thin as proved by the case study presented in this paper. Calandra [2] suggested that their shape took inspiration from ephemeral and mobile structures such as Persian and in general Oriental canopies: some of them became a new and widespread typology of domes (ribbed vaults), others, perhaps more fragile and complex to construct had no luck and probably few existing examples can be found at Hadrian's Villa or in Baia (Campania).

2. AIMS OF THE STUDY

The ongoing research on Hadrian's Villa vaulted spaces is aimed at understanding, through the use of accurate mesh models from active sensors, the ancient designing techniques used by architects for developing novel shapes in accord to Hadrian's new aesthetic.

In particular, the study carried out on Small Thermal Baths (Figure 1) wants to underline common features of its octagonal hall with other buildings' parts like the southern hall of the Piazza d'Oro and the reverse-curve pavilion of the Accademia.



Figure 1: View of the interiors of Small Thermal Baths, Hadrian's Villa.

This small mixtilinear room presents strict formal relation with the other two building plans, but compared with them the first presents a good preservation state of the original vault.

For this reason, the hints coming from its geometric-constructive algorithm may lead to a more accurate reconstruction of the other two halls, as well as to a more accurate virtual reconstruction hypothesis of the vault in comparison to those achieved by Verdiani et alii [14] by means of a previous laser scanner survey carried out in 2007 and 2008 using phase shift technology devices.

The purpose of this study is then to achieve a 3D reconstruction using a uniform integrated framework applicable to other buildings, not only related to Roman Imperial archaeological remains, since it integrates several research fields.

In particular, it is useful to underline the importance of a constant dialogue among institutions in charge of management and conservation of the site, experts in the field of laser scanner technologies, specialists in the interpretation of ancient architecture.

This last aspect of the research is crucial since it provides a "cultural guidance" for the achievement of a 3D reconstruction whose main purpose is to show and underline the difference between these models of vaults in comparison to the others belonging to the archaeological site of the Villa.

Trying to go beyond the shared certainty expressed in previous studies on the influence of Hadrian's Villa on later architectures, from Late Antiquity [7] to Baroque [11], the analysis carried out on the survey of this octagonal hall intends to make new hypothesis. It concerns the possibility of a common geometrical construction for the blending of different 2D shapes forming the lower and the upper section of a dome: in other words, in which way they blended the polygonal profile of the impost into the upper oculus or spherical cap.

By this point of view the problem did not start with Hadrian's "pumpkins" or tents since

it was present also in previous examples and the octagonal hall of the Domus Aurea (1st century AD), at a later stage followed by the decagonal nymphaeum called Temple of Minerva Medica (4th century AD).

3. METHODOLOGY

The study carried out on the dome of the octagonal hall of Hadrian's Villa small thermal baths is the product of an integrated methodology based on a double "filter": the first cultural, the latter technical.

The first aspect considers 3D accurate representations as a sort of data storage in which several facts are encoded: first of all every building was the outcome of different steps (a methodology) used by ancient designers to solve practical issues.

Architects' work has always been the outcome of a process which begins taking into consideration specific requirements (functional, aesthetic, economical, etc.). Practical issues are then solved by putting into practice a series of theoretical regulations, efficiently expressed by Vitruvius in the first book of his treatise.

The second aspect deals with a true and full usage of reverse modelling (RM) applications as tools for the investigation of ancient architectural generation. This concept is quite far from being widespread among experts of computer applications in the field of archaeology and Cultural Heritage, since RM software is underrated and relegated, in the majority of cases, to the role of mesh processing tool.

The investigation on octagonal hall is then a perfect case study for testing this approach, being sufficiently preserved, but at the same time, also being the product of a sudden construction; it presents some irregular features (distorted shapes and not parallel walls).

In accordance with Vitruvius, the first step of the investigation is to split the built object in different design steps: this aspect consists in the use of reverse modelling applications for the achievement of two kinds of drawings, which corresponded in the ancient times to two as-

pects of the so called *Dispositio*. This term is used by Vitruvius for describing three aspects of the designing process: *Ichnographia*, *Orthographia* and *Scaenographia*. They all correspond to different steps of designing, drawing and planning the building.

Ichnographia and *orthographia* in particular can be described as the "art of designing the plan" and the "art of designing the section (or facade)". These two apparently confused translations underline the importance of avoiding trivial translations that make them correspond to current terms belonging to architect's technical lexicon: *ichnographia* equal to plan, *orthographia* equal to section.

For instance, when studying the design of an ancient theatre starting from its survey it is important to distinguish between the plan and the geometrical construction (a sort of graphic algorithm) used by the architect for the dimensioning of the scenic building, the stage, the way to split the bleachers of the cavea, etc.

In this sense when having at disposal the 3D mesh of a room, of a dome, etc., it is crucial the way cutting planes are set into the digital environment, both horizontally and vertically.

These planes are not to be used for the achievement of a correct, but impersonal, set of drawing and sections: this planes have to correspond to the sections designed (not just drawn) by the architect and possibly should be the base for an heuristic phase in which the researcher tries to understand how ancient architects applied their mathematical and geometrical knowledge for solving practical problems.

Other two aspects have to be taken into account during the investigation process based on drawings extracted from highly detailed mesh models from active sensors: they deal with other two words used by the Latin author that imply the concept of modularity and proportion. They are *Simmetria* and *Eurhythmia*: the first term refers to the necessity of grid-based design; the second is the use of proportions in order to achieve a harmonic composition of all the elements characterizing the parts of a building.

Proportions are achieved through a modular grid that facilitates the composition process and the calculations needed for the distribution of all the areas and lengths that will give form to the final construction, forming part of its aesthetic and fulfilling specific requirements. In other words, for ancient architects it was necessary to find a proportion between different parts. For instance a portico is formed by a roofed area and a void surface in the middle: those two areas should comply with the requirements of the customer meanwhile the architect has to manage those quantities through his geometrical mathematical and typological knowledge. For this reason it is quite common to find simple relations (ratio), as 2:1, 3:5, $1:\sqrt{2}$, between lengths, areas, etc. [8].

The module on which *simmetria* is based is the product of the addition of a convenient number of standard measures; it is a three-dimensional modular grid (not just limited to the plan) aimed at giving common proportions to the whole building and its parts. A module is an abstract concept and should not be misunderstood or confused with ancient anthropometric system (standard measures) in use during Roman Empire: a module, for example, could be a quantity equal to 3 *pedes* (feet) or 4 *cubiti* (arms), but in general it doesn't correspond to a standard measure as a *pes* (foot) or to a *cubitus* (arm), but it is a specific "quantity" that will be present in each part of a construction.

These examples provide a general framework, a sort of cultural "filter", for the achievement of an adequate awareness when huge amounts of data are at disposal. In this sense it is important to underline the importance of a coherent strategy for the study of such a complex matter, since many practical problems emerged at the moment of using so dense and accurate models from laser scanner and SFM applications. Abundance of data corresponds to more problems to manage and solve, for this reason is important to achieve common standards and terminology for data

interpretation. In this general framework, Reverse Modelling (RM) applications play an important role even if they were developed in different professional fields with different purposes. Such tools, if correctly used, match perfectly with the objectives of data interpretation explained in this paper from three different points of view:

- RM applications provide an interactive and efficient set of commands that facilitate the process of definition of reference planes, vectors and points necessary for the achievement of 2D templates;
- RM applications distinguish between two meanings and aims of 2D sections: documentation and interpretation. In one case they provide sections, contour lines, polylines directly extracted from the mesh model; in the second case they use 2D templates as reference materials for the redrawing process of specific geometric features.
- A last tool is then the possibility to compare idealized models and shapes based on NURBS, solids, etc. with the corresponding original mesh, giving to the whole reconstruction and interpretation process more accuracy and reliability.

The convenience of RM in the investigation of vaulted spaces is then evident since roofing always had more complex geometrical and constructive structures in comparison with the remains of simple masonry walls, where geometrical investigation can only concerns *ichnographia*.

4. HYPOTHESIS ON THE DESIGN PROCESS OF THE OCTAGONAL HALL

The survey of the Small Baths was carried out in the context of a scientific collaboration among three Italian universities: Florence, Pavia and Bologna (Figure 2). The octagonal room shows significant portions of the original roofing well-preserved, but to the naked eye it is very difficult to grasp its morphological genesis since it is a unique and somewhat pioneering. As stated by Dr. Elena Calandra, So-

printendenza per i Beni Archeologici del Lazio, many of the vaulted spaces of Villa Adriana, particularly those characterized by ribbed vaults, seem they have taken inspiration from ephemeral architecture of eastern origin [2] and also this last preserved dome could be the result of this influence.

The octagonal hall was acquired using a phase shift laser scanner device Zoller + Fröhlich Imager® 5010. After the registration process carried out with Leica Cyclone, the point cloud model was exported using the .PTX format toward INUS Technology Rapidform XOR 3, where every scan was cleaned and then integrated into two mesh models: one for the interiors and the other for the roof oriented East.

The highly detailed mesh obtained from laser scanner survey was used as template for the production of a series of sections obtained thanks to RM software.



Figure 2: Surveying activities at the octagonal hall during a workshop focused on Small

Thermal Baths, 22-27 April, 2014.

Due to the intrinsic lack of accuracy of the construction which produced deformations and deviations between the orientations of walls, two kinds of sections have been used:

- the first set corresponds to a few numbers of horizontal and vertical planes in relation with the assumptions expressed in the previous paragraph (*ichnographia*, *orthographia*, etc.);
- the second set of sections consist of contour lines since they provide the hints for the understanding of the general flow of curvature of the complex and deteriorated shape of the vault's remains.

In figure 3 a general plan and three sections represent the current state of the hall. CC and DD section cut the octagonal through the middle of the four bended walls of the mixtilinear octagon, while AA section slices the model northwards along two rectilinear walls.

A first statement concerns the CC and DD sections because the consistency of the dome's remains allows a reliable achievement of two best fitting circles: thanks to semi-automatic re-drawing tools it was possible to establish a first hypothesis about requirements and constraints to which the ancient architect had to answer. These two circles of about seventeen *pedes* have the centre at the same height of the lower edge of the four windows illuminating the hall and in addition they are tangent to the height of the original floor.

The AA section seems to underline the presence of a linear slope in correspondence to the lunette similar to a longitudinal ridge rib. The four best fitting circles present slightly different radii ranging from 3,762 to 4,367 m. Three of them are quite similar, while the one toward the circular hall of the *laconicum* is far bigger than the others.

Among the researches focusing on the mixtilinear halls of Hadrian's Villa and the implication of grid-based design for the development of their *ichnographia*, Rakob [12] and Ytterberg [15] provide interesting suggestions, con-

firmed by the measurements extracted from the highly detailed model: the basic module is equal to 5 *pedes* and it was used by the ancient architect for the construction of the octagonal shape (Figure 4) starting from a square whose edge lengths are equal to 35 *pedes* and diagonal equal to 50 *pedes*. For the design of the curved sides it seems more fitting a different solution,

since the one proposed from Rakob diverges from the best fitting circles more than the one proposed in this paper.

Rakob suggests a simple and smart geometrical construction based on the linear edge of the octagon equal to 15 *pedes*, but the average based on best fitting circles is equal to 13, 5 *pedes* (Figure 4).

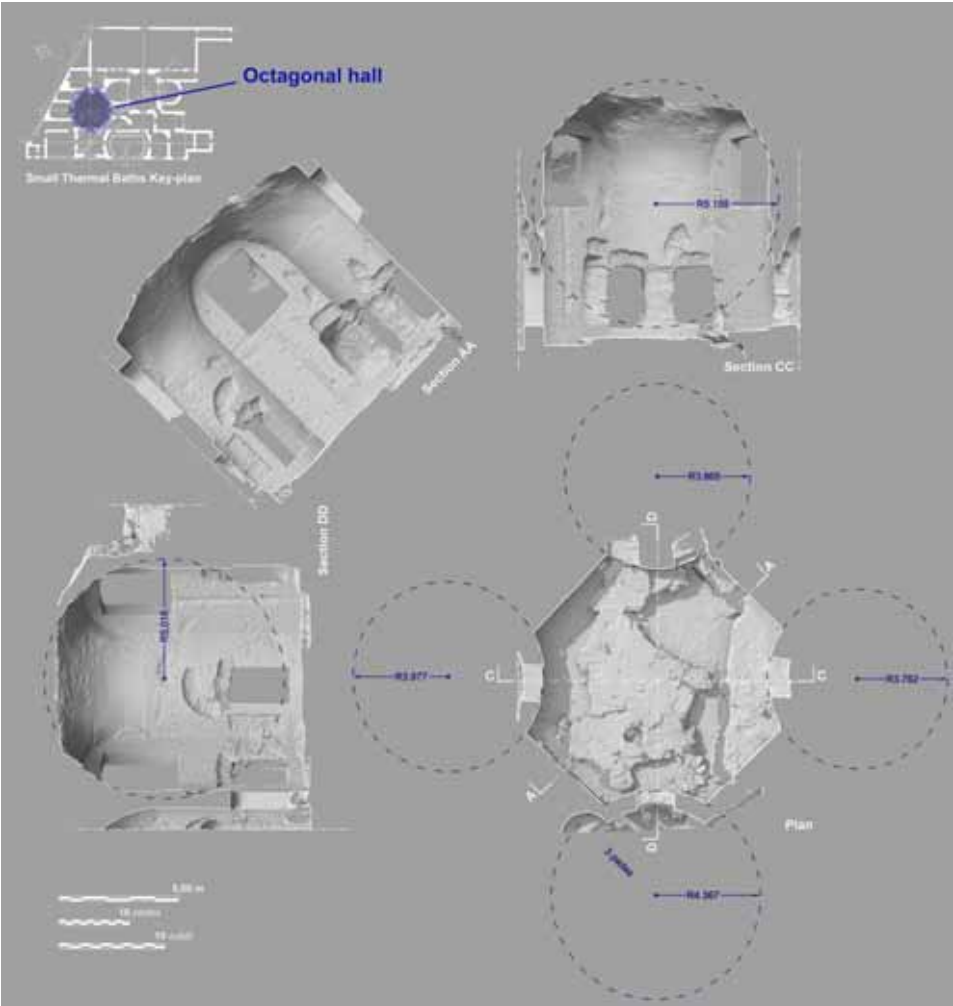


Figure 3: plan and main sections of the octagonal hall of Small Thermal Baths.

The geometrical construction shown in this paper uses another approach that takes into consideration the possibility that ancient architects might have used a first modular structure, after followed by a series of inscribed circles and rotating squares, in some way similar to the one adopted for the design of theatres. This hypothesis shown in figure 4 C consists in connecting the intersection of curved sides with the liner edges of the octagon in a way that produces 2 squares reciprocally rotated of almost 45 degrees. The tangent circle to the edge of these squares has a radius equal to 13.46 *pedes* which corresponds to a better approximation of the average circumferences in comparison to Rakob's proposal. Moving the investigation toward the *orthographia* of the octagonal hall is possible to notice other interesting aspect of its design: in figure 5 there is a comparison between the vertical measure the octagonal hall, achieved through the best fitting circle and the horizontal general grid that produced the *ichnographia* of the same room. The circle inscribed into the section of the hall presents a diameter equal to the side of the fundamental square on which is based the plan. *Ichnographia* and *orthographia* are deeply linked to the modular structure adopted by the architect and, as other buildings of the same epoch e.g. the Pantheon, intending to achieve *eurhythmia* through explicit modularity and simple ratios. Even more interesting is the result of these assumptions along AA section. Aiming at keeping this second *orthographia* compatible with the other one (sections CC and DD) through the achievement of 35 *pedes* of height, it is possible to notice how the same two squares used for the bended sides can provide further help for the interpretation of ancient design. The linear slope underlined in AA section fits with the angle formed by the edges of these inscribed and rotated squares. Even in this case the designing of both plan and section are linked by some relation: in the first case by the modular structure, in the latter the geometrical construction is based on circles and squares.

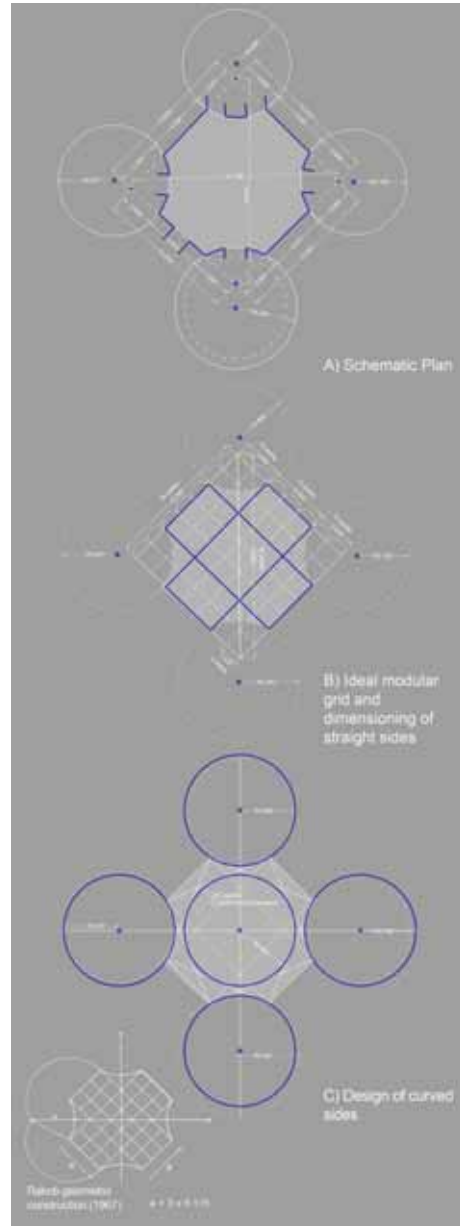


Figure 4: dimensioning and designing hypothesis of the plan.

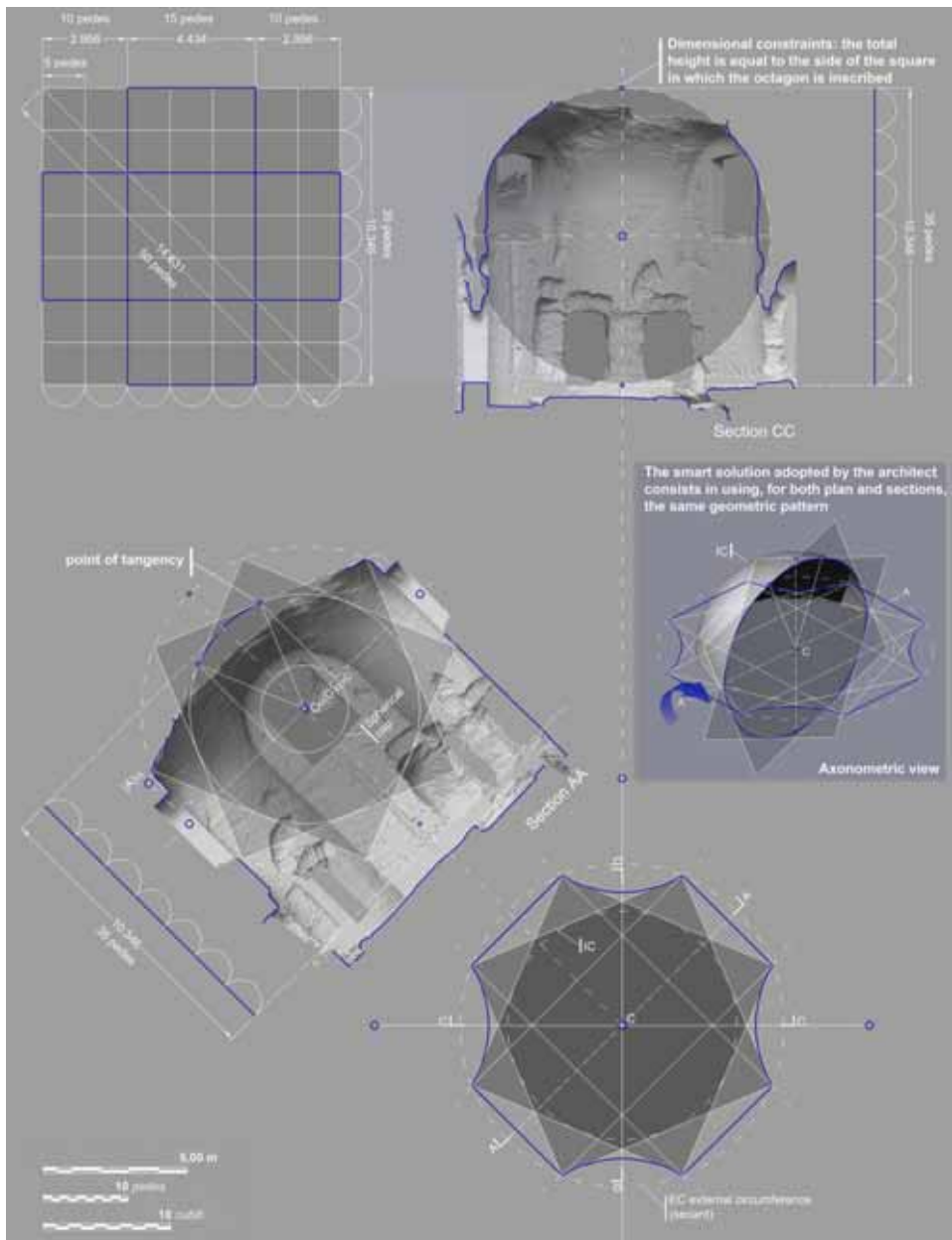


Figure 5: dimensioning and designing hypothesis of the main sections.

So it seems that the architects may have used the same geometric pattern for plan and elevation: in this way, thanks to the tangency between the inner circle (IC) and the squares it was simple to obtain the spherical cap. In the axonometric view in figure 6 is outlined the overturning process, but in any case, other questions are still present: how did they blend the mixtilinear shape of the octagon into a spherical cap? The usage of inscribed squares seems to suggest that ancient architects wanted to appeal to some kind of exhaustion method for solving this transformation, but at this phase of the research is too difficult to make a reasonable hypothesis.

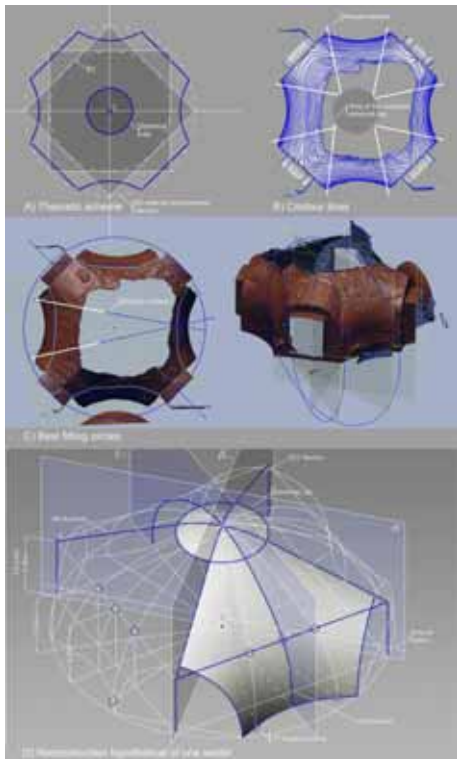


Figure 6: curve network construction.

The problem is even more complicated due to the state of conservation of the vault and by the lack of accuracy in the construction. It's a pity since the geometry beneath its design was so advanced and smart. The solution adopted for giving a partial not comprehensive solution comes from contour lines obtained from the highly detailed mesh: in figure 6 A other two squares are inscribed between the external and the inner circle. The sides of these geometric shapes can be identified in the general flow of horizontal contour lines. Focusing the attention on contour lines (figure 6 B) it is possible to grasp that the geometric continuity of the vault is no more than a skilful fake, since the blending from a convex wall to a the concave spherical cap on the top of the dome is achieved through the use of simple ruled surfaces, whose vertices lie on eight "invisible" planar sections that approximately converge on the outer border of the spherical cap.

In figure 6 C the procedures are pointed out inside RM applications for the achievement of best fitting circles that then will be used inside a NURBS modelling application for the construction of one sector of the vault intrados.

Another aspect of this reconstruction hypothesis is shown in figure 6 D where the blue colour represents the straight horizontal line where all the three curves, that blend a shape into another, have to pass. This line is obtained by displacing one of the squares in figure 6 A approximately at the height of the lunette.

5. CONCLUSIONS

The final 3D model of the vault was obtained using both NURBS and Subdivision Surface modelling techniques. The first model corresponds to one sector of the intrados; it is an idealized representation based on a regular grid and for this reason a comparison between the 3D model from laser scanner and this theoretical representation was not a priority of this research. Despite that, the NURBS-based repre-

sensation provides to this model a not perfect curvature flow since ancient architects had confidence in the finishing phase, in which the concrete vault was covered in abundant plaster (almost ten centimetres or even more). This solution could hide the not perfect conti-

nuity of the surface and for this reason Sub-D modelling techniques was used for the achievement of the final model (figura 7) since it provides, thanks to Catmull-Clark subdivision rule, a better curvature of the original idealized model.

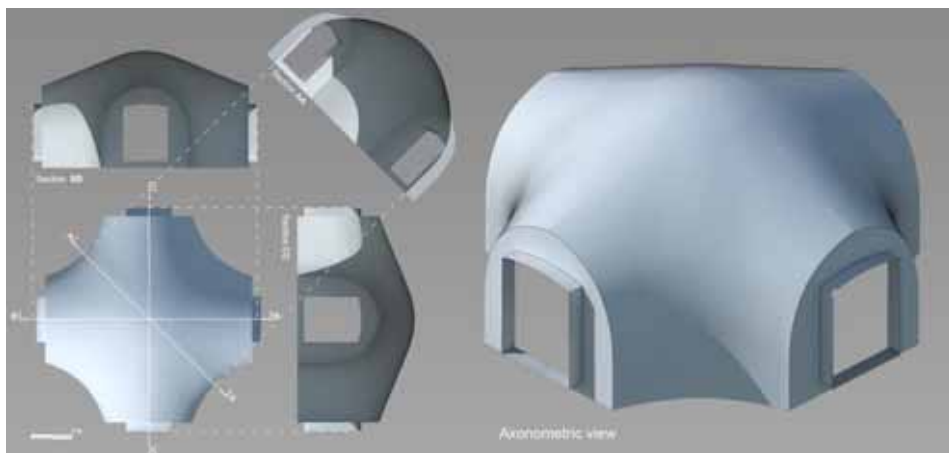


Figure 7: reconstruction hypothesis of the hall's intrados.

The next phase of this study will deal with the survey of another building belonging to the Hadrian's Villa, the Reverse-Curve pavilion of the Accademia. A hypothetical reconstruction of the southern hall of the Piazza d'Oro was recently proposed by Adembri at alii [1] developing the hypothesis by Hansen [9] who guessed the possibility of a great cupola covering the southern Hall of that part of the Villa. The results of this research, in terms of methodologies and knowledge on ancient designing techniques for vaulted spaces, should be applied to this last hall with the aim of having at disposal a complete panorama of these tent-like cupolas. In addition, another investigation line should start with the aim of deepening geometrical aspects concerning the problem of exhaustion and squaring of the circle from geometers and mathematicians as Antiphon and Bryson of Heraclea.

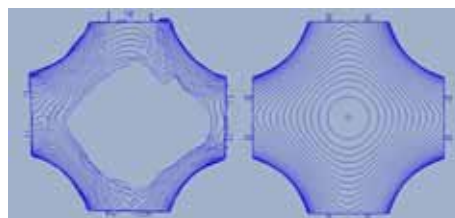


Figure 8: comparison between controur lines (left: current state, right: hypothesis).

ACKNOWLEDGMENTS

This paper is part of an ongoing surveying project carried out by the University of Florence (Professor S. Bertocci), University of Pavia (Professor S. Parrinello) and the University of Bologna (Professors L. Cipriani and F. Fantini). The surveying campaign on the Small Thermal Baths (April 22-27, 2014) was carried out in collaboration with Ph.D. Sergio Di Tondo from Microgeo S.R.L. (Campi Bisenzio, Florence).

The authors would like to thank Dr. Elena Calandra and Dr. Benedetta Adembri of the Soprintendenza per i Beni Archeologici del Lazio, for the collaboration and the suggestions provided for the development of this paper.

REFERENCES

- [1] B. Adembri, S. Di Tondo, F. Fantini, and F. Ristori. Casi di studio a partire da rilevamenti laser scanner a Villa Adriana: interdisciplinarietà ed integrazione dei metodi. In E. Calandra, G. Ghini, and Z. Mari (edited by), *Lazio e Sabina. Lavori e Studi della Soprintendenza per i Beni Archeologici del Lazio*, 10, pages 283-286. Roma, 2014.
- [2] E. Calandra. Adriano princeps e committente. In *Forma Urbis, numero monografico su Villa Adriana*, XVIII (8), September 2013.
- [3] P. F. Caliari. *Tractatus logico sintattico. La forma trasparente di Villa Adriana*. Edizioni Quasar, Roma, 2012.
- [4] E. Cary. *Dio Cassius Roman History*, Volume VIII, Books 61-70, Loeb, Harvard, 1925.
- [5] G. Cinque. Analisi geometriche e progettuali in alcuni complessi di Villa Adriana. In *Romula*, 9: 62-63, 2010.
- [6] S. Di Tondo. *La forma di Villa Adriana nel territorio Tiburtino* (PhD thesis). University of Florence, Faculty of Architecture, Survey and Representation of Architecture and Environment. Section Architecture and Drawing, Florence, December 2007.
- [7] S. Di Tondo and F. Fantini. Gli ambienti a pianta centrale di Villa Adriana: anticipazione di una nuova estetica. In *DISEGNARECON*, (9), June 2012.
- [8] F. Fantini. The use of reality based models for the interpretation of ancient architecture: experiences of reverse modelling at Masada. In S. Bertocci, S. Parrinello, and R. Vital (edited by). *Masada notebooks, Report of the research project 2013*. Edifir, Firenze, 2013.
- [9] E. Hansen, J. Nielsen, J. Asserbo, and T. Jespersen. Due cupole a Villa Adriana. Calcoli statici. In *Analecta Romana Instituti Danici*, XXXV-XXXVI (app. 1). Munksgaard, Copenhagen, 2011.
- [10] T. Opper. *Hadrian Empire and conflict*. The British Museum Press, London, 2008.
- [11] W. L. MacDonald and J. A. Pinto. *La costruzione e il mito da Adriano a Louis Khan*. Electa, Milano, 1997.
- [12] F. L. Rakob. *Die Piazza d'Oro in der Villa Hadriana bei Tivoli* (PhD thesis.), Munchen, 1967.
- [13] E. Salza Prina Ricotti. *Villa Adriana. Il sogno di un imperatore*. L'Erma di Bretschneider, Roma, 2001.
- [14] G. Verdiani, M. Pucci, and A. Blanco. The Small Baths in Hadrian's Villa. A ground test for enhancing the approach to the digital survey and reconstruction for archaeologists and architects. In *14th International Congress Cultural Heritage and New Technologies*. Vienna, 2009.
- [15] M. R. Ytterberg. The Hidden Order of Hadrian's Villa, and the Order of Modern Architecture. In *Nexus Network Journal*, 15 (1): 127-154, 2013.

ABOUT THE AUTHORS

1. Luca Cipriani is Associate Professor at Alma Mater Studiorum, University of Bologna (Italy). His research is mainly focused on the systematization of digital techniques for classical architecture by means of 3D digital surveying of monuments and Cultural Heritage.
2. Filippo Fantini is Researcher at Alma Mater Studiorum, University of Bologna (Italy) and Ph.D. in *Science of Surveying and Representation of Architecture and Environment*. Since 2004 he has been carrying out a research and survey on Hadrian's Villa.

GEOMETRY OF GOTHIC VAULTS BASED ON HISTORICAL SOURCES.

Anna KULIG

Cracow University of Technology, Poland

ABSTRACT: New technologies applied in the modern world became a very useful tool for researchers of historical architecture. The author made an attempt at a virtual reconstruction of the spatial form of the vault. Virtual reconstruction was created based on drawing preserved from the fifteenth century in church. Analysis of a drawing carved on the plaster covering the Wall of Gothic church In Szydłowiec resulted in unexpected conclusions. The author writes about this unique discovery, explaining the secrets of techniques of Gothic bricklayers. The author is informing that the drawing was discovered in the Gothic church of St. Sigismund during conservation work conducted between 1970 and 1992 by Professor Zofia Medwecka (Academy of Fine Arts in Cracow). The unknown drawing on the plaster was preserved; it was a huge architectural plan carved on the aisle wall. It turned out to be design of the vault built in this church over the presbytery. A full-scale diagram (13x8,5m) was used by carpenters and bricklayers for the processing and assembly of profiled stone arches creating the massive ribs of the stellar net vault. The scale of the diagram and its precision are unique: it became an inspiration for new interpretations of the workshop of former stonemasons. Using computer technology a projection of the span was made with the drawing of an eight-branch star; profiles and curves for ribs were applied using only data from the engraved diagram (without spatial measurements). The virtual model was subsequently compared to the original in the facility: both visually and with respect to measurements. The architectural design on the wall along with the executed facility which has been preserved for five centuries has high scientific value and its modern interpretation in computer simulation is aimed at making historical sources about old constructions technologies and the role of geometry in the stonemason's workshop more available

Keywords: Geometry, Gothic vaults,

1. PRELIMINARY

Gothic rib vaults preserved in Polish historical buildings of sacral character were discussed. Examples of complex geometrical shapes, i.e. of rich and decorative patterns, were selected. They were compared with the oldest available drawing representations originating from the 16th and the 17th century. The analyses focused on design and realization of vaults as well as the role of the geometry in the workshop of builders.

2. INTRODUCTION

Rib and stellar vaults erected on the terri-

tories of Poland since the beginning of the 14th century constitute one of the biggest accomplishments of the Gothic architecture. Constructions and spatial forms of Gothic vaults were developed in the Western and Central Europe as a result of experimenting and constant perfecting of the construction and the aesthetics.

Numerous publications focused on vaults, their origins, and development of their forms, inspirations and styles [11, 13]. In turn the methods of the medieval builders, their workshop, work organization and the vault erection process were considered only rarely and to insufficient extent. One of the obstacles here is

lack of historical records from the times when these constructions were erected. What have remained are merely bills, contracts and records of sponsors. The actual constructors have remained anonymous. Poland has no equivalent of the Villard's sketchbook. Hence the original structures of buildings remain the main source in the studies. Moreover, the craftsmen traditions of those times dictated strict professional confidentiality and learning was taking place in a workshop of a master builder through a direct contact.

Hence, an exceptional value, considering the acute lack of historical plans and designs, represents thus an enormous carved out architectural plan from the turn of the 15th and 16th centuries, randomly discovered (at the end of the 20th century) in the church in Szydłowiec. Behind layers of plaster it remained hidden for five centuries.

3. DESIGN AND REALIZATION OF A VAULT BASED ON THE EXAMPLE OF THE GOTHIC CHURCH IN SZYDŁOWIEC.

The built of bricks church of St. Sigismund in Szydłowiec near Radom (approximately 130 km south of Warsaw) was erected in 15th century. It has a traditional form: a wide nave with side chapels and a narrower and lower presbytery. Sandstone walls of irregular thread with corner strokes indicate a medieval origin. The interior of the nave is covered with a wooden ceiling while the presbytery, sacristy and chapels have rib vaultings. The most ornamental, late Gothic vaulting has been preserved over the presbytery.

On the northern wall of the nave an enormous drawing (of dimensions 13x8.5 m – Figure 1) representing a full-scale plan – design – of this vault is present. This drawing by a medieval architect was discovered in the 20th century (by prof. Medwecka who between 1970 and 1992 was carrying out preservation works of the interior of the church The drawings and the photographs were made in 1991 by the em-

ployees of the Faculty of Architecture of the Warsaw University of Technology (using traditional method, directly, on scaffoldings). Drawings in the 1:10 scale are available in the archives of the Polish Architecture Department of the said Faculty (the drawing has the dimensions of 848 x 1270cm). The size and shape of the drawing match the dimensions of the presbytery. The drawing was carved out in plaster.



Figure 1: Szydłowiec. Church of St. Sigismund. Nave. View of the northern wall with the carved out drawing dated at the turn of the 15th and 16th century.

The carvings are approximately 2 mm depth and were made legible using dark putty. The drawing consists of vertical, horizontal and skewed at 45° lines of equal weight.



Figure 2: Szydłowiec. Church interior. View of the northern wall. Vault plan is marked in red color.

Certain portions of the plan are currently covered or damaged through adaptations that were made to the interior (holes, railings). Nevertheless these portions can be easily reconstructed in the documentation thanks to the regularities of the form and analogies between the arches.

The Gothic drawing contains information on the positioning of all ribs, their intersections, support sites and the depth of the embedment in the wall. The width of the three constantly parallel lines corresponds with the rib's width (16 cm – Figure 2). Aside to the horizontal projection, in the lower part of the wall, a projection of the arch – rib – as well as its profile (marked as a semi-section – in other words as a superimposed section) were situated.



Figure 3: Szydłowiec. View of the presbytery with the Gothic vaulting of the turn of the 15th and 16th century.

The lower portion of the drawing cuts short as a consequence of later adaptations, i.e. raising of the floor level. The carving was carefully surveyed, recorded and preserved [2,4]. Interpretations of the phases of creation of this design, in the form of a two dimensional plan basing on a square of a side of dimension equal to the width of the presbytery, were also undertaken.

However a spatial reconstruction of the structure did not follow back then – it was carried out thus recently and is presented hereinafter.

The interior of the presbytery is covered by a rib vaulting realized according to the plan preserved in the close vicinity. The vaulting is divided into three identical orthogonal arches and an eastern multilateral termination. Considering the soffit, this type of ceiling qualifies to sail vaults with lunettes, while when one considers the pattern that is formed by the ribs the vault is classified to stellar-network vaults (Figure 3). Made of stone, profiled and massive ribs flow onto the walls strengthened in these sites by external buttresses.

Each of the arches is divided by ribs into a dozen or so small compartments filled with liernes. The composition is dominated by the shapes of large centrally situated eight-pointed stars, the arms of which however do not reach the walls. The subsequent three stars composed of the arch segments of the ribs create a chain like sequence thanks to the overlapping rhomboidal arms situated along the axis of the vault.

As a consequence the borders between the arches are hidden which results in a unification of the entire arrangement. Compartments between the ribs form sail like hulls; these are actually not segments of a sphere even though they resemble such. Segments of a sphere are formed only around the keystones.

The interior is intimate and low-lying (approximately 9 m). As a consequence, very good conditions exist as far as observation of the entire arrangement and its details is concerned (Figure 4). The regularity of the arrangement,

its carefully selected proportions and painstakingly worked out details are noticeable. A contrasting color scheme visible in brown/brick-red polychromies of the ribs and white colored compartments further improves the aesthetics of the interior.



Figure 4: Szydłowiec. Vaulting of the presbytery, a panoramic view in the direction of the southern wall.

4. MODERN, AUTHOR'S INTERPRETATION OF THE GOTHIC DRAWING

In this section, authors can acknowledge all collaborators (people or companies), funding agencies, etc. who contributed in any form to the paper or the research reported on it. This section is not numbered.

The carved out plan is an absolutely precise horizontal projection of the Gothic vault erected at the end of the Gothic. Nowadays new possibilities for an interpretation of this historical document arise. The author attempted to

reconstruct the geometrical structure basing primarily on this plan (Figure 5). A geometrical model of the vault was created in its "perfect" form, according to the historical drawings, as if repeating the former builders (Figure 6). The source materials for the reconstruction were the actual historical materials, i.e.:

- a) the carving and its survey made by the students of the Warsaw University of Technology, including the reading of the radius of the base arch,
- b) nowadays observations and photographs of the vault and the carved out plan,
- c) photographs of similar drawings of Gothic vaults preserved in museums. A parchment drawing of 1468, preserved in the archives in Vienna, in Akademie der Bildenden Künste in Wien, under the Inventory number 17069, representing a vault of identical structure as the vault in Szydłowiec but in a smaller scale proved to be of particular value [4, 10].

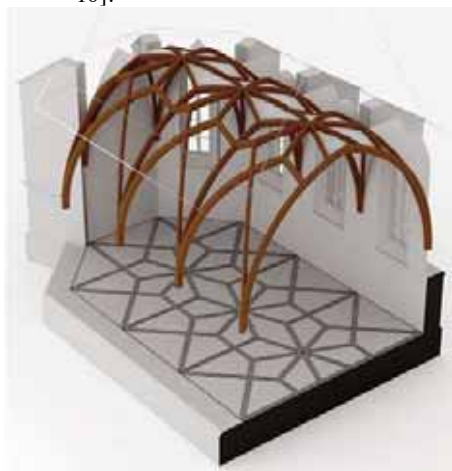


Figure 5: Virtual model. Axonometric projection of the presbytery of the St. Sigismund's church. Reproduced structure of the vault over the "bed" with marked out plan copied from the wall carving.

And what data were missing to create a

computer reconstruction of a geometrical model of the vault? In particular altimeter measurements were not denoted, as only very rarely they were added to old plans. One can assume hence that the builders managed this aspect differently, without arithmetic, using exclusively geometrical methods and basic geometrical figures, i.e. the square, triangle and circle.



Figure 6: View of the stellar-network vault. Rendering of the interior.

In this particular case (as per author's assumption) the regular octahedron of which three sides are enclosed by the eastern wall of the presbytery was the starting figure (Figure 7). This enclosure is visible at the top of the drawing. An octahedron is formed by superimposing two squares rotated with respect each other by 45° (i.e. by rotation of squares).

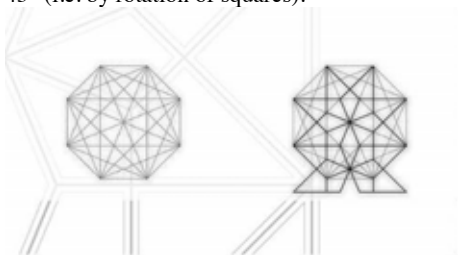


Figure 7: Phases of creation of the vault plan. Octahedron is the initial figure.

Octahedrons led the geometrical composition in each of the arches. Drawing diagonals and marking their intersections provides already an outline of an eight-pointed star. Having all diagonals left was considered superfluous as would result in too small compartments and hence certain sections were eliminated. Divisions with big quadrangular compartments were retained and the shape of a star was exposed. The middle portion of the octahedron with the drawing of the star became a repeating module for the three arches (Figure 8).

In order to reproduce the elevations of the characteristic intersections between the ribs, the carving of a rib's arc fragment and the reading of its radius (equaling 5.5 m) were referenced. In this case the elevation of the keystones was determined over this semicircle. The diameter of the semicircle (i.e. 11 m) was compared with the sum of the lengths of sections of a single arch drawn between the long walls of the presbytery (section AB, BC, CD in Figure 9). The two values were identical. This allowed concluding that the semicircle of a wooden spiral reinforcement that constituted a scaffold during assembly and bonding of stone rib segments was "broken" twice at an angle and was situated between the walls as a screen. This presumption corresponds with the description and marking of semicircle points on a drawing that is stored in the archives of Vienna and which represents, although in a smaller scale, a plan and a cross-section of a vault of identical shape and size [10]. This kind of a drawing was most likely used by builders of the vault.

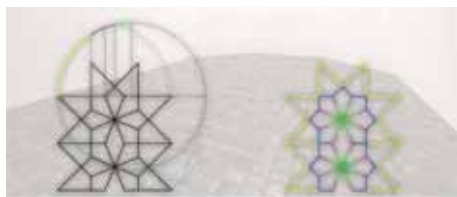


Figure 8: Phases of creation of the vault plan. Octahedron is the initial figure.

A vault of the same type was erected earlier in Germany – in 1498 – in Owingen near Überlingen. The vault of the presbytery has also numerous analogies in European works – in Germany (in Stammheim of 1487 and Ziegelheim of 1518), in Czech Republic (of approximately 1511 in Horní Dvůr and of 1485-1490 in the Prague Castle in the Vladislav Hall), as well as in Austria (in Schluderns of approximately 1500) [12,13].

The flat and linear character of the drawing originates from its practical application as a technical instruction during the construction phase [5,8].

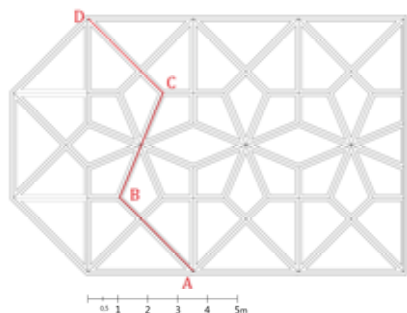


Figure 9: Plan of the vault based on the Gothic drawing. Sections AB, BC, CD indicate the polygonal chain used to position the semi-circular spiral reinforcement of 11 m diameter.

Architectural drawing of the Gothic played a practical role, was a method of work and a stage in preparations and planning. It was often created directly on the construction site as a “drawing to be lost”. Typically it was prepared on a drawing floor levelled with a layer of clay or plaster, less frequently on walls. It never played aesthetic or artistic functions and hence originates its modest form, linear character as well as lack of spatial representations, illusions, colors, shadings etc. Drawings of this kind represented predominantly projections, views and details. Their characteristic feature was precision and accuracy. Construction drawings (on

paper or parchment), as written sources confirm, constituted a part of construction contracts and care was taken to ensure their preservation. They were considered common assets by guilds as well as a scholar material and an inspiration source. Drawings were copied, modified, used as patterns and carriers of ideas. The concept of copyrights was unknown; contrary it was often downright suggested to base upon an existing famous piece and freely use its plans. Architectural drawings of Gothic times served over long construction periods, for instance during erections of cathedrals which lasted several generations. Preserved parchment drawings of particularly big dimensions (as the drawing of the façade of the cathedral in Strasbourg of 1383 of 4m length) could have played a role of a prospect, a representation of the future work, and displayed to the public encouraged founders and donators. Drawings were inherited by the subsequent site managers who continued the work. Transition of traditions and skills was taking place through the drawing as well. In guilds lessons of drawing were carried out based on re-drawing of existing works and creating new variants and patterns. One has to emphasize the universal values of Gothic drawings, i.e. the role of the presentation of the technical achievements of the time as well as the role in the transfer of the tradition.

5. VAULT VISUALIZATION ON THE BASIS OF THE HISTORICAL PLAN USING DIGITAL TECHNOLOGY.

A scaled version of the plan was created using computer technology based on survey data supplemented with reconstruction data for the missing rib fragments. Relevant points were marked on the plan and the intersections between the ribs were assigned appropriate elevations. Matching elevations were marked with the same color. A plan matching the one carved out on the wall was created but this time it contained additional information on the spatial positioning of the points over the plan. A semi-circle of the base circle of 11 m diameter served

as the initial figure for the assignment of the elevations (the semicircle was situated in each of the arches not diagonally but slantwise along the polygonal chain which is marked on the drawing as AB, BC, CD). “Poles” of appropriate heights were positioned on the plan (the highest in the keystone and lower ones moving away from it). These poles supported then arcs of circles of constant radius arranged according to the plan (Figure 10). The geometrical structure was then enclosed and arcs of the ribs were given thickness and a profile. Compartments between the ribs were filled with ruled surfaces. The model was then viewed by positioning a camera in a position similar to that of a spectator inside the original presbytery. A rendering was prepared using graphical and light effects (Figure 6).

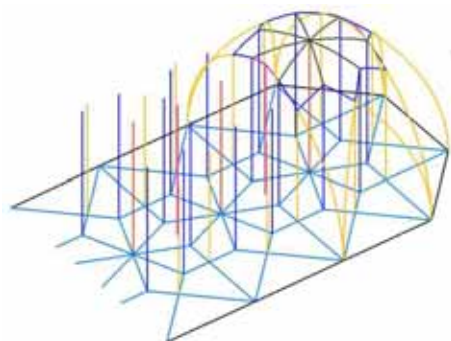


Figure 10: Stage in creation of the virtual model. Positioning of “poles” supporting the fragments of the arcs of the circles of constant radius but varying positions of their centers.

The resulting virtual image was almost identical with a photograph of the existing object and the spatial character of the vault’s structure, its clarity and appeal were even better visible as they were not obscured by the elements of the subsequent interior design. Computer assisted reconstruction of the vault’s structure, application of profiles and curvatures to the ribs, all were made based exclusively on the data available from the carving. The virtual model was

then compared to the original object – visually and by means of measurements (in key points correctness of the rule applied for the determination of the elevation of the relevant points, i.e. intersections between the ribs, was verified). The visualization confirmed strict relationship between the carved out drawing and the realized vault.

6. CONCLUSIONS

Old architectural plans were preserved only partially, even though their existence is acknowledged. A situation where both a historical plan and a based upon it historical realization coexist (as in the church in Szydłowiec) is very rare.

The plan carved out on the wall and the realized work represent together high scientific value. Modern interpretation through a computer simulation is intended to popularize historical sources on old building methods and the role of the geometry in the workshop of old builders.

The virtual model made by the author in 2013 is a new interpretation of not only the drawing itself but also of the principle of the construction of the spatial form – in accordance with the 15th century drawing.

ACKNOWLEDGMENTS

The research has been funded by the National Science Centre under the research grant number 2011/01/D/HS2/02288.

REFERENCES

- [1] Broniewski T., *Pomiary i zdjęcia architektoniczne*, Katowice 1949.
- [2] Brykowska M., *Metody pomiarów i badań zabytków architektury*, Warsaw 2003.
- [3] Brykowska M., *Architektura kościoła św. Zygmunta w Szydłowcu w świetle najnowszych badań*, [in] *Z dziejów parafii szydłowieckiej*, editor J. Wijaczka, Szydłowiec 1998, pp. 145-164.
- [4] Brykowska M., Kunkel R., *Sprawozdanie z praktyk inwentaryzacyjnych Instytutu Hi-*

- storii Architektury i Sztuki w 1991 r.*, KAiU XXXVII, 1992, book 3, pp. 297-308.
- [5] Frazik J.T., *Małopolskie paralele późnogotyckich sklepień w wiedeńskim zbiorze średniowiecznych rysunków architektonicznych*, TKUiA, vol. XII, 1978, pp. 227-237.
 - [6] Frazik J.T., *Analiza materiału, techniki i stratygrafii murów, jako metoda badawcza dzieł architektury zabytkowej*, BHS XXXI, no. 1, 1969, pp. 121-123.
 - [7] Frazik J.T., *Ze studiów nad warsztatem architekta i budowniczych w średniowieczu*, Cracow University of Technology, Monograph 102, Krakow 1990.
 - [8] Frazik J.T., *Technika średniowiecznych rysunków architektonicznych*, TKUiA XXVII, 1995, pp. 195-203.
 - [9] Gruszecki A., *Projekt a ostateczny kształt detalu architektonicznego w XVI-XVII wieku*, [in] *Podług nieba i zwyczaju polskiego. Studia z historii architektury, sztuki i kultury ofiarowane Adamowi Miłobędzkiemu*, Warsaw 1984, pp. 253-257.
 - [10] Koepp H. *Die gotischen Planrisse*, Wien 1969, Abb.458. Umzeichnung
 - [11] Łodyńska-Kosińska M., *Geometria architektów gotyckich*, KAiU IX, 1964, book 2, pp. 89-114
 - [12] Müller W., *Grundlagen gotischer Bautechnik*, München 1990, pp. 32, 73, 84, 155.
 - [13] Nussbaum N., *Das gotische Gewölbe: eine Geschichte seine Form und Konstruktion*, Darmstadt 1999
 - [14] Rewski Z., *Majstersztyki krakowskiego cechu murarzy i kamieniarzy XVI-XIX wieku*, Wrocław 1954

ABOUT THE AUTHORS

Anna Kulig, PhD of Cracow University of Technology

ON THE GEOMETRY OF PSEUDO-KOSSEL AND BREMSSTRAHLUNG INTERFERENCES

Stefan ENGHARDT¹, Jürgen BAUCH¹ and Frank HENSCHL²

¹ Technische Universität Dresden, Germany ² Dresden University of Applied Sciences, Germany

ABSTRACT: A new real- structure analysis method involving hard X-rays (Bremsstrahlung) is being developed for the study of mono- and polycrystalline materials. This allows the simultaneous visualization of a spatial X-ray shadow microscopic image (radiographic information) with a spatial diffraction pattern (crystallographic information with possible lattice defects). Future developments will focus on the correlation between inspected and disturbed sample volumes and reflection distortions (by reconstruction methods). This would permit the interpretation of defects on a crystallographic scale along with the radiographic information.

In order to reach this goal it is necessary to study further the detectable interference curves. Starting from a point source (inside an anti-cathode), the X-rays spread as a spherical wave. This wave is diffracted at planes of the crystal lattice in accordance with Bragg's law. This physical equation describes in a geometric way the condition for which a particular angle of incidence leads to constructive interference. For each single angle and set of parallel lattice planes this condition is fulfilled for one curve of second order on the (idealized) plane sample surface as conic section of the circular diffraction cone. The so-called Pseudo-Kossel-technique considers the interference patterns in the direction of the reflected wave. The maxima of these waves lie on quadratic cones which are visible as fourth order curves on a detector plane. According to the previous investigations of Borrmann in 1949 the transmitting X-rays spread into two components per wave after leaving the sample. The two associated directions are coplanar and form the same angle with the normal vector of the lattice plane. The resulting algebraic curves on a planar detector are then of up to eighth order.

Keywords: conic section, algebraic curves, ruled surface, X-ray diffraction, Kossel-technique.

1. INTRODUCTION

X-ray diffraction is used in various material science analysis techniques. These are, in particular, applied to the study of crystalline materials since they make up over 95% of solid matter. The various diffraction methods lead to lines and points on a suitable X-ray detector. The study of these geometric structures enables conclusions to be drawn about the properties and defects of the researched sample.

A novel approach has been developed which combines the creation of an X-ray diffraction image along with a X-ray shadow microscopic image used within the field of classical radio-

graphic testing and movement of the detector plane around one fixed sample [1–2]. The aim of this method is to allow for the identification of sample properties on the macroscopic and micro- (or nano-) scopic scale. Figure 1 shows one image, out of a series, which utilized this method with the use of hard X-ray Bremsstrahlung (spectrum of X-rays up to relatively high energies). In Figure 1 it is easy to make out the outline of the cuboid sample with two cylindrical drill holes. Additionally, there are lines of higher brightness (intensity) which consist of a fan-like structure and are related to fog-like effects in their close proximity (cp. Fig. 6).

In order to study the structure and defects of

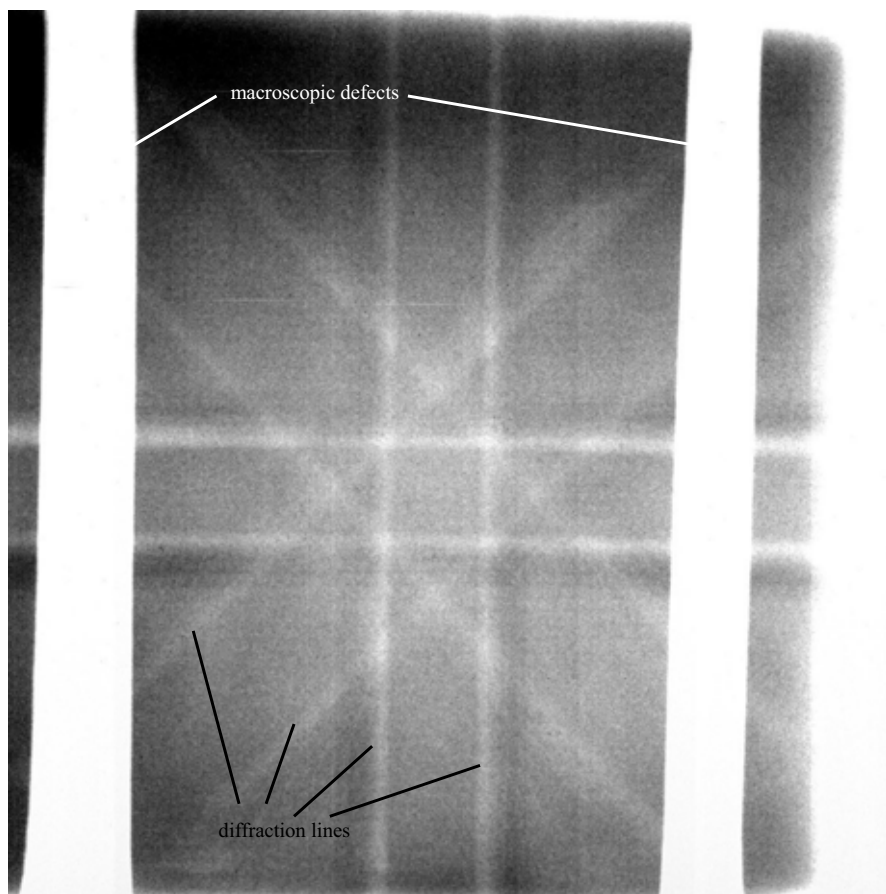


Figure 1: Image (digital detector) of an $[100]$ orientated 2 mm thick copper single crystal, with two drill holes (macroscopic defect) and diffraction lines (crystallographic defect).

the material it is necessary to first study the resulting lines and their geometry. Subsequently, the well understood central projection may then be used in conjunction with the information contained in the diffracted X-rays to get additional or more precise information as compared to the classical individual methods of diffraction or radiographic testing.

2. GENERATION OF X-RAY DIFFRACTION IMAGES

Most solid matter is of crystalline structure. This can be mathematically characterized by three-dimensional lattices where the lattice points correspond with the positions of atoms or ions (translation symmetry). It is possible to find families of parallel planes such that the lattice points lie on the planes of one such fam-

ily. In the field of diffraction one commonly used way of describing these sets of planes are the so-called Laue indices (hkl) . In the case of cubic symmetry this triple can be viewed as the normal vector $(h\ k\ l)^T$ of the corresponding planes in a specific Cartesian coordinate system. Von Laue discovered the principle of X-ray diffraction in crystals. This was honored with the Nobel Prize in 1914 [6]. W.L. and W.H. Bragg simplified these laws to one equation:

$$\sin \vartheta_{hkl} = \frac{\lambda}{2d_{hkl}} \quad (1)$$

This was also awarded with a Nobel Prize in the following year 1915 (cp. [4]). In Equation (1) the physical quantity λ is the wavelength of the incident X-rays. The parallel lattice planes (hkl) have the separation d_{hkl} which may change for example under the influence of heat or physical deformation and result in residual stresses. According to Equation (1) there is only one incidence angle ϑ_{hkl} of the incident X-rays onto the planes (hkl) which leads to constructive interference. In this case the X-rays are diffracted and reflected with the identical angle ϑ_{hkl} . The resulting set of interference fringes is detectable on a suitable screen.

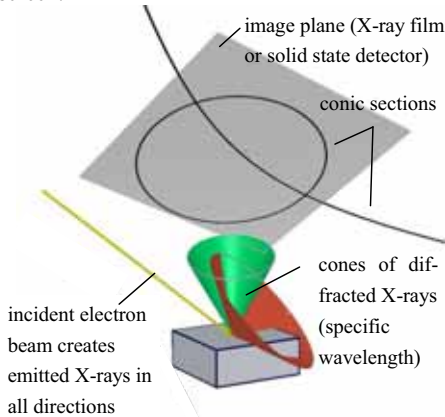


Figure 2: Outline of the Kossel method with cuboid sample.

Consider a point source of X-rays inside the crystal which emits spherical waves in every direction. In this case Equation (1) is satisfied for each lattice plane by a straight circular cone of X-rays (lattice source interferences). A plane detector screen thus records conic sections. This effect is central to the Kossel method which uses an electron beam to create the X-ray source inside the crystal. This method leads to highly precise information about the examined crystal. The corresponding diffraction images may be interpreted for instance by using focal curves [5]. Fig. 2 gives an overview of the three-dimensional situation and resulting curves.

2.1 Pseudo-Kossel-Method

The Kossel method has one significant disadvantage: the wavelength of the generated X-rays depends on the material being examined. For instance the characteristic wavelength of silicon (Si- K_α) is too large to cause interference effects inside the crystal. Consequently, the method can not be used for the most important microelectronics materials. To avoid this problem it is necessary to use a different wavelength.

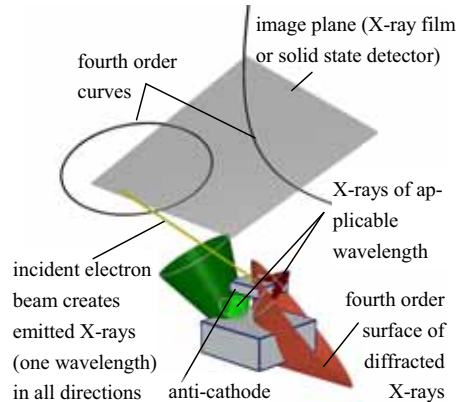


Figure 3: Outline of the Pseudo-Kossel method with cuboid sample and anti-cathode above it.

One possible way to reach this aim is by the

insertion of an anti-cathode of different material above the sample into the incident electron beam (cp. Fig. 3) which leads to the so called Pseudo-Kossel method. Fig. 4 shows an image of gallium arsenide which was generated by using X-rays with the wavelength of copper. The resulting reflections are generally not curves of second order (cp. section 3.1). The Pseudo-Kossel technique allows for a precise determination of crystallographic orientation and lattice constants [7].

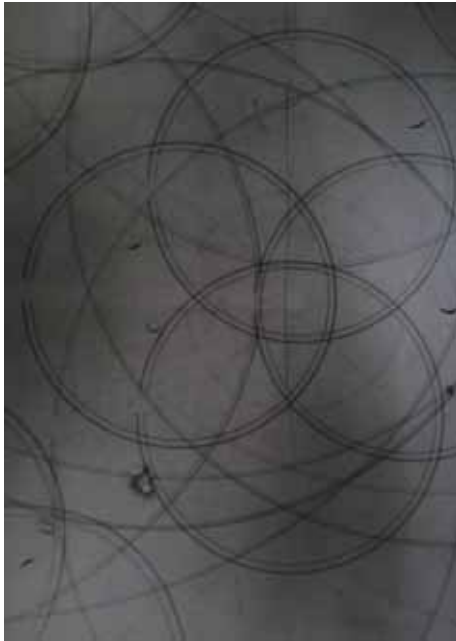


Figure 4: Pseudo-Kossel image of GaAs with Cu anti-cathode and synchrotron excitation [9].

2.2 3-D Bremsstrahlung Diffraction with Radiographic Testing

This technique utilizes divergent X-ray Bremsstrahlung, which is a continuous spectrum of X-rays with a range of energies that correlate with relatively short wavelengths when compared to the Pseudo-Kossel technique. Higher

energies allow the X-rays to transmit through the sample material. Depending on the type of material and its thickness the transmitted and attenuated X-rays can be detected on an appropriately placed screen. The central projection of the sample is then visible on the image plane (cp. Fig. 1, 6). A rotation of the sample allows for additional information to be obtained, for example the dimensions of a defect inside the material [1-2].

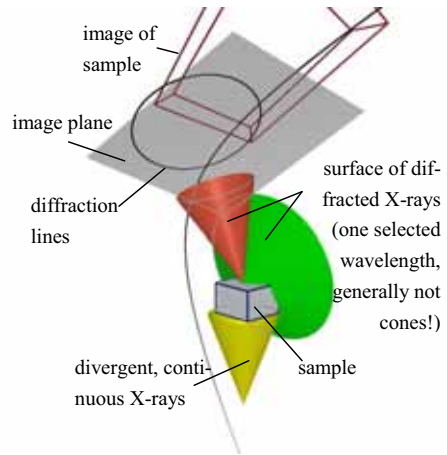


Figure 5: Outline of 3-D Bremsstrahlung diffraction method, including the central projection of the sample.

Images made with this technique show lines which are not tied to the macroscopic measurement of the sample (compare lines in Fig. 1 and 6). The observed differences in intensity, when compared the shadow of the sample, can be explained by X-ray diffraction processes similar to those found in the Pseudo-Kossel method explained in section 2.1. Due to the short wavelengths in the range of λ_{\min} the pattern appear as “Kikuchi-like”.

The continuum of generated wavelengths leads in accordance with Equation (1) to a continuum of concentric cones per lattice plane. Consequently, the visible lines are not as sharp as those created by the Pseudo-Kossel tech-

nique. The Bremsstrahlung X-rays consist of waves with different wavelengths. These waves are of different intensity which leads to a difference of brightness in the image.

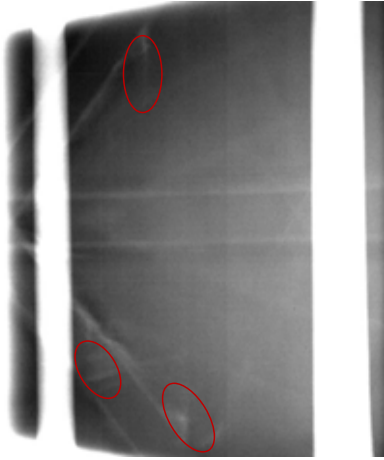


Figure 6: 3-D Bremsstrahlung diffraction image with sample rotated by 208° compared to Fig. 1. The red ellipses mark the fog-like structures (caused by crystallographic defects) that are hardly visible in Fig. 1.

3. GEOMETRIC MODEL OF DIFFRACTED X-RAYS

Both the Pseudo-Kossel and Bremsstrahlung diffraction techniques utilize divergent X-rays. In order to simplify the following considerations, only X-rays of one fixed wavelength are considered. In the case of Bremsstrahlung one may select the X-ray wavelength with the largest intensity. Equation (1) gives only general information about the influence of lattice planes on X-rays. The paths of the X-rays inside the crystal follow laws which are studied within the field of the dynamic theory of diffraction. One result of this theory is the Borrmann effect [3]. The geometric interpretation of this effect implies the movement of diffracted X-rays inside the sample, on lattice planes. Thus every incident beam which satisfies Equation (1) follows its projection onto the lattice plane (hkl) .

In the following let S with position vector \vec{s} be the source of the divergent X-rays. Fig. 7 shows a cut through the sample, one lattice plane through L_1 and L_2 with normal \vec{n}_{hkl} and two X-ray waves originating at S that satisfy Equation (1). This condition is fulfilled for one specific cone of X-ray beams per lattice plane and wavelength. The cone has axis direction \vec{n}_{hkl} , apex S and half opening angle $90^\circ - \vartheta_{hkl}$. Consequently, these X-ray waves intersect with a plane sample surface in a conic section (through A_1 and A_2 in Fig. 7).

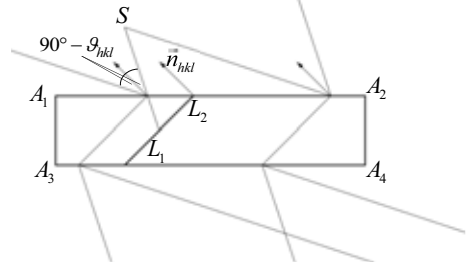


Figure 7: Incident and diffracted X-rays (according to Equation (1) and Borrmann effect).

3.1 Reflected X-Rays

Consider the conic section of X-rays satisfying Equation (1) with the plane sample surface. The diffraction of X-rays that do not transmit the sample volume can be viewed as reflection of the X-rays on the lattice planes (hkl) at the sample surface (cp. Fig. 7). The reflected X-rays can thus be calculated as reflections of the incident beams on the straight line in direction \vec{n}_{hkl} through the intersection with the surface. Thus the surface of reflected X-rays is a ruled surface. The reflection across a line through the origin O can be described by a matrix $L \in \mathbb{R}^{3 \times 3}$ which is the rotation matrix with axis direction \vec{n}_{hkl} and angle 180° . If \vec{n} is a unit vector parallel to \vec{n}_{hkl} , then:

$$L = \begin{pmatrix} 2n_1^2 - 1 & 2n_1n_2 & 2n_1n_3 \\ 2n_1n_2 & 2n_2^2 - 1 & 2n_2n_3 \\ 2n_1n_3 & 2n_2n_3 & 2n_3^2 - 1 \end{pmatrix} \quad (2)$$

Let $\vec{c}(u)$ be one parameterization of the conic section on the sample surface. Then the image of the X-ray source S is

$$\vec{s}'(u) = L(\vec{s} - \vec{c}(u)) + \vec{c}(u) \quad (3)$$

Thus, for every point on the directrix $\vec{c}(u)$ there is one known point on the ruling: $\vec{s}'(u)$. Therefore the director curve is known and the ruled surface RS_1 has the following form:

$$RS_1 : \vec{r}(u, v) = \vec{c}(u) + vL(\vec{s} - \vec{c}(u)) \quad (4)$$

Since both directrix and director curve are of second order the ruled surface is a quartic surface. Special cases include surfaces of lower order. For instance $\vec{c}(u)$ may be a circle and L be the identity. In this case RS_1 describes a cone. Fig. 8 shows an example of this surface where the directrix is a hyperbola. One alternative approach of describing the Pseudo-Kossel diffraction as wide angle diffraction can be found in [8,10].

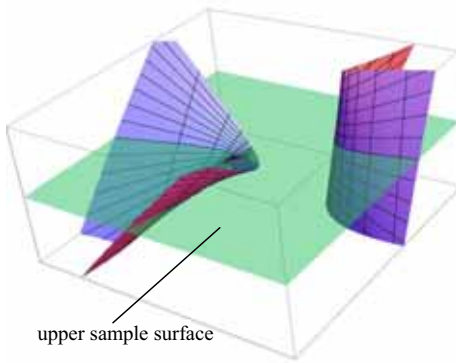


Figure 8: Example of the Pseudo-Kossel surface RS_1 (fourth order) with hyperbolic directrix. The intersection of surfaces RS_1 with the image plane (reflections) results in curves of fourth order.

With the parameterization in Equation (4) it is easy to generalize the results for different sample surfaces. Figure 9 shows the resulting surface of a cylindrical sample that is of the same material as that in Figure 8. For general surfaces this may lead to curves of arbitrary order.

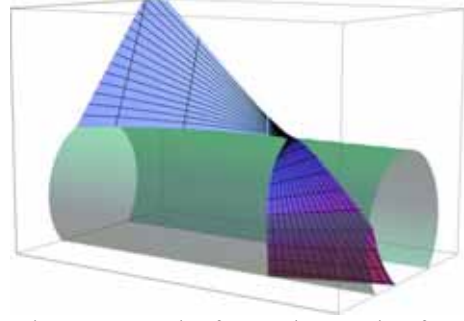


Figure 9: Example of a Pseudo-Kossel surface on a cylindrical sample (cp. with Fig. 8).

3.2 Transmitted X-Rays

The transmission of diffracted X-rays (according to Equation (1)) through the crystalline sample happens on the lattice planes [3]. Thus it is important to look at the X-rays while they are transmitted. Since the beams follow straight lines through the crystal (cp. Fig. 7) they form one ruled surface with directrix $\vec{c}(u)$ from section 3.1. The parallel projection of the incident rays onto the lattice plane (hkl) leads to the director curve. Generally this curve is of second order.

Let the ruled surface be RS_2 and P be the matrix of the projection. Consequently, the surface may be described as:

$$RS_2 : \vec{r}(u, v) = \vec{c}(u) + vP(\vec{c}(u) - \vec{s}) \quad (5)$$

Analogue to section 3.1, RS_2 is a surface of up to fourth order (cp. Fig. 12).

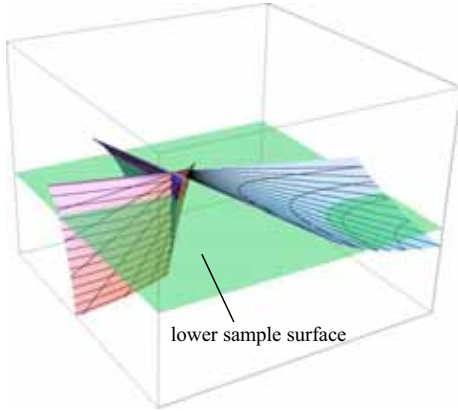


Figure 10: Ruled surface RS_3 of X-rays after leaving the crystal – cp. Fig. 11.

To understand the principle of Bremsstrahlung diffraction, consider simplified sample parameters. Therefore, let the sample surfaces be planes (may be achieved by sample preparation). In general the surface RS_2 then intersects with a sample surface along a curve of fourth order (for $v \neq 0$). By following the X-ray paths shown in Fig. 7 two additional ruled surfaces can be drawn. These are detectable on a screen:

$$RS_3 : \vec{r}(u, v) = \vec{d}(u) + v(\vec{c} - \vec{s}) \quad (6)$$

$$RS_4 : \vec{r}'(u, v) = \vec{d}(u) - vL(\vec{c} - \vec{s})$$

L is the rotation according to Equation (2). RS_3 and RS_4 are generally of up to eighth order (cp. Fig. 10-12).

In the following passage the central ideas of the proof are presented:

Let \vec{d} and \vec{c} be plane curves. When using an appropriate coordinate system RS_3 may be written in the following way ($\varrho_{ij} \in \mathbb{R}$):

$$\begin{aligned} r_1 &= d_1 + v(\varrho_{10} + \varrho_{11} c_1 + \varrho_{12} c_2) \\ r_2 &= d_2 + v(\varrho_{20} + \varrho_{21} c_1 + \varrho_{22} c_2) \\ r_3 &= v(\varrho_{30} + \varrho_{31} c_1 + \varrho_{32} c_2) \end{aligned} \quad (7)$$

Let d be of fourth order and c be of second order. Consequently, there exists a representation of these curves as

$$0 = \sum_{i,j}^{i+j \leq 4} \delta_{ij} d_1^i d_2^j \quad (8)$$

and likewise

$$0 = \sum_{i,j}^{i+j \leq 2} \varsigma_{ij} c_1^i c_2^j. \quad (9)$$

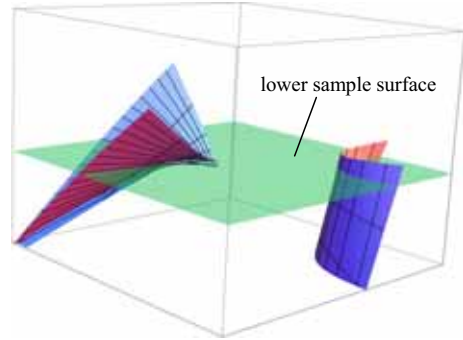


Figure 11: Ruled surface RS_4 of X-rays after leaving the crystal – cp. Fig. 10.

Equation (8) leads along with Equation (7) to:

$$0 = \sum_{i,j}^{i+j \leq 4} \delta_{ij} (r_1 - v(\varrho_{10} + \varrho_{11} c_1 + \varrho_{12} c_2))^i \cdot (r_2 - v(\varrho_{20} + \varrho_{21} c_1 + \varrho_{22} c_2))^j \quad (10)$$

Since Equation (10) is true for every v , every coefficient of v^n has to be zero for every n . This is true especially for $n=4$. There are curves d with δ_{ij} such that the coefficient of v^4 is of fourth order in c_i where the coefficients of c_1^n are zero for every n but $\varsigma_{10}, \varsigma_{20} \neq 0$. Since the ruled surface exists there has to be a solution. Consequently, polynomial addends of fourth order in c_i have to be appear in the term. This is only possible by raising the order of r via the introduction of terms like r_3^8 (cp. Equation (7)). The resulting sur-

face is thus of up to eighth order.

Special cases of r as a curve of lower than eighth order are possible, especially depending on the ruled surface RS_2 . For example, consider \vec{n}_{hkl} normal to both the upper and lower sample surface. Then the resulting ruled surfaces are cones.

4. CONCLUSIONS

The analysis of the novel method which combines Bremsstrahlung diffraction and central projection of the sample has led to the description of the diffracted X-rays as a ruled surface. This surface is detectable on a plane detector as curve of up to eighth order. Depending on the model of the photon path inside the sample (RS_2 , Borrmann effect) the surface may be of lower order. This has to be verified by experiments.

It has been shown that there is a direct correspondence between those surfaces and the surface of fourth order which is used by the Pseudo-Kossel method. This allows for the simulation of the corresponding X-ray diffraction images and provides a basis for a high accuracy determination of lattice parameters and crystal orientation.

Additionally it has been shown that a generalization of the Pseudo-Kossel lines for arbitrary sample surfaces is feasible.

The interpretation of the ruled surfaces leads to possible explanations of effects like the "fog" observed in Fig. 6 as intersection of the ruled X-ray surface of fourth order inside the crystal with the drill holes of the sample.

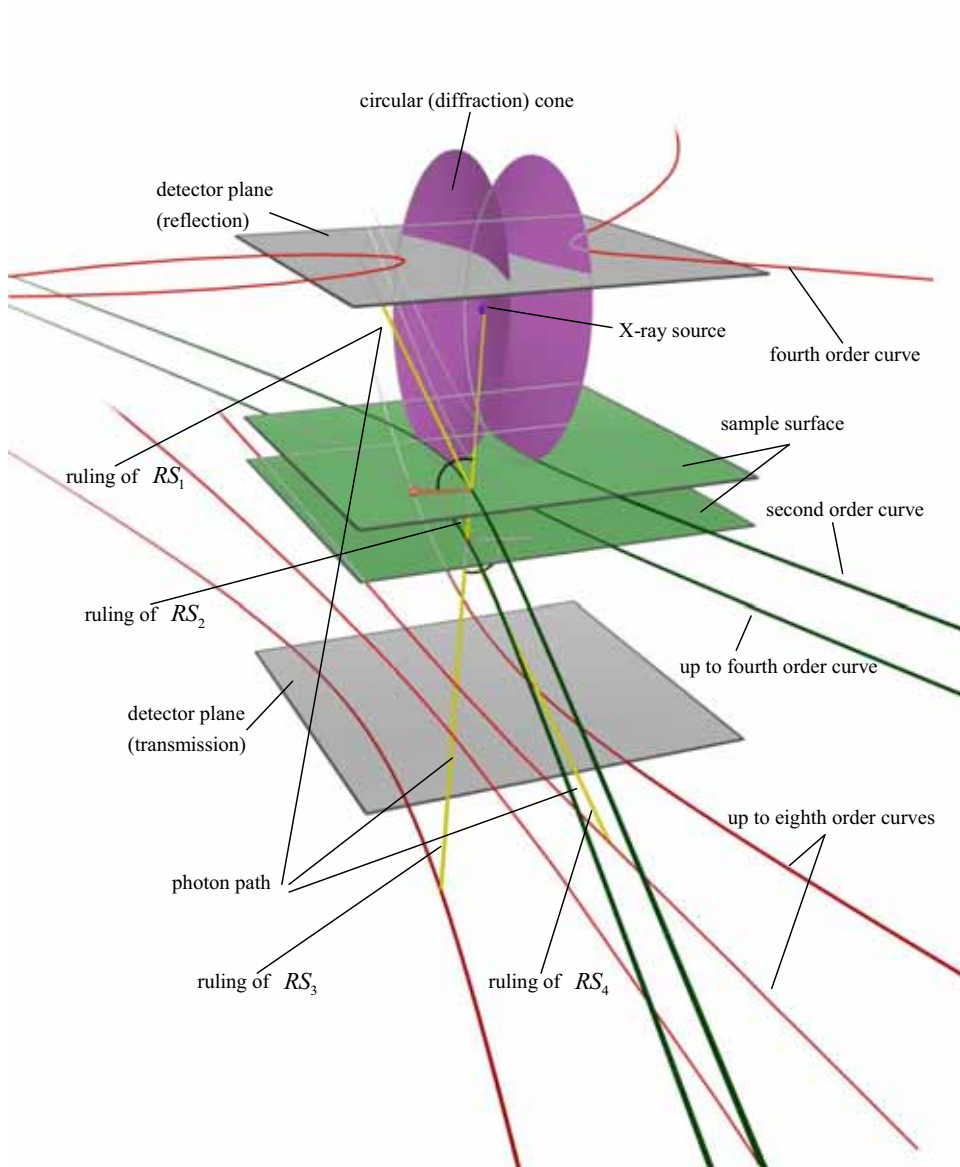


Figure 12: Overview of resulting curves and transmission paths via Pseudo-Kossel and Bremsstrahlung diffraction techniques.

ACKNOWLEDGMENTS

We are thankful to Dr. R. Boucher and Prof. Dr. M. Hamann for their help.

REFERENCES

- [1] J. Bauch, M. Böhling, D.C. Lupascu, H.J. Ullrich. Verfahren und Vorrichtung zur Registrierung von Realstruktur-Informationen in massiven Kristallkörpern mittels Röntgenstrahlung. Patent, Patent DE 10 2008 008 829.
- [2] J. Bauch, F. Henschel, D. Wünsche and H.-J. Ullrich. Verfahren und Einrichtung zur dreidimensional orts aufgelösten, bildhaften Darstellung von Informationen zur kristallographischen Realstruktur und zur makroskopischen Defektstruktur in massiven Kristallkörpern. Patent, WO 2011/003408 A1.
- [3] G. Borrmann. Die Absorption der Röntgenstrahlung im Fall der Interferenz. *Zeitschrift für Physik*, 127 (4): 297–323, 1950.
- [4] W. L. Bragg. Nobel Lecture: The Diffraction of X-Rays by Crystals.
http://www.nobelprize.org/nobel_prizes/physics/laureates/1915/wl-bragg-lecture.html, 6 2014
- [5] F. Henschel and J. Bauch. Application of Focal Curves for X-ray Microdiffraction methods. *Journal for Geometry and Graphics*, 17 (2): 205–211, 2013.
- [6] M. v. Laue. Nobel Lecture: Concerning the Detection of X-ray Interferences.
http://www.nobelprize.org/nobel_prizes/physics/laureates/1914/laue-lecture.html, 6 2014
- [7] E. Langer. Der Einfluß von Kristallfehlern auf Kossel- und Weitwinkel-Interferenzen. Dissertation. TU Dresden, 2004.
- [8] H. Lin. Methodische Beiträge zur Kossel- und Pseudokossel- Technik. Dissertation. TU Dresden, 2000.
- [9] H. Lin, H.-J. Ullrich, J. Bauch and S. Keßler. The Detection of Pseudo-Kossel-Lines Excited by Synchrotron Radiation on Si- and GaAs-Wafers. *HASYLAB Annual Report 2000 (1)*.
- [10] A. Uhlig. Beiträge zur Auswertung röntgenographischer Beugungsaufnahmen. Dissertation. TU Dresden, 1984.

ABOUT THE AUTHORS

1. Stefan Enghardt, Dipl.-Math., is a research assistant at the Institute for Material Science and member of the working group “physical material diagnostics” at TU Dresden. He can be reached by e-mail: stefan.enghardt@tu-dresden.de or through the postal address: TU Dresden, Institute for Material Science, 01062 Dresden, Germany

2. Jürgen Bauch, Prof. Dr.-Ing. habil., is a professor at the Institute for Material Science at the TU Dresden and head of the working group “physical material diagnostics” at TU Dresden. He can be reached by e-mail: juergen.bauch@tu-dresden.de or through the postal address: TU Dresden, Institute for Material Science, 01062 Dresden, Germany

3. Frank Henschel, Dipl.-Math. is a laboratory engineer at the Faculty of Informatics/Mathematics at the University of Applied Science Dresden. He can be reached by e-mail: henschel@informatik.htw-dresden.de or through the postal address: HTW Dresden, 01069 Dresden.

THE GEOMETRY OF THE “PROSPECT GEOMETRIQUE” BY MICHELI DU CREST (1754) A QUANTITATIVE ANALYSIS WITHIN A HYBRID LEAST-SQUARES ADJUSTMENT FRAMEWORK

Helmut KAGER¹, Martin RICKENBACHER², and Andreas RONCAT¹

¹Research Groups Photogrammetry and Remote Sensing,

Department of Geodesy and Geoinformation, Vienna University of Technology, Austria

²Swiss Federal Office of Topography *swisstopo*, Wabern, Switzerland

ABSTRACT: The “Prospect Geometrique” by Jacques-Barthélemy Micheli du Crest (1690–1766) is the oldest panorama visualization of the Alps with a scientific background. Its creator was a remarkable figure in Swiss history: He studied physics, astronomy, geodesy, and cartography in France and invented a temperature scale which was in use until the 19th century. On the other side, his political activities led to a sentence to death in absence, later transformed into a life sentence which he spent mostly in Aarburg castle (canton of Aargau, Switzerland). During his imprisonment, he developed a concept of land survey and realized the most possible in his situation with primitive instruments: the *Prospect Geometrique*, i.e. a central perspective surveying of his field of view mapped onto a circular cylinder with vertical axis. This technique was once considered as *photogrammetry without photographs* by the renowned Swiss cartographer E. Imhof.

The term *Prospect* used by Micheli means the correct geometric representation of the landscape from the observer’s viewpoint. A first numerical investigation of its “correctness” was conducted in 1995 [10] and later refined in 2012 [11]. While the referenced publications have a more cartographic scope, this paper is intended as a sequel to the latter, focusing on the geometry and stochastics for analysing the Prospect thoroughly in a hybrid adjustment approach by least squares. Several input categories for the adjustment have been considered: (a) the upper and lower scale bars for azimuth and (b) left and right scale bars for elevation given on the outlines of the Prospect (see facsimile in Figure 4) and (c) control points identifiable in contemporary official geodetic maps. The core issue of this study was to determine how exact Micheli could construct the Prospect, based upon the determination of the actual projection centre, the calibration of the radius and orientation of the projection cylinder and further a potential exaggeration of the elevation scale.

Keywords: Photogrammetry, Central Perspective, Map Projection, Adjustment by Least Squares

1. INTRODUCTION

Panoramas are considered as map-related representations in the field of cartography so that geometric aspects are inherent to this kind of representation. This study is dedicated to the “Prospect Geometrique” by Jacques-Barthélemy Micheli du Crest (1690–1766). It is the oldest panorama of the alps with a scientific background.

By choosing the term “Prospect Geometrique”, its result is supposed to reflect correctly the actual geometric situation of the landscape from the observer’s viewpoint in a central projection. The outdated french term “Prospect” has its roots in the latin term “prospectus” (for view) and means the view of a distanced landscape. In our case, the observer’s viewpoint was a very special location:

Micheli Du Crest was imprisoned in Aarburg castle when he worked on his Prospect.

The aim of this paper is to assess the geometric quality of and the accuracy of its representation in the Prospect. In contrast to a previous study [10], a purely digital data flow is used in this analysis by means of GIS (Geographic Information Systems) software and a photogrammetric bundle block adjustment tool.

Section 2 presents the Prospect and its creator. In Section 3, the workflow and analysis strategy are presented. Results are given in Section 4, whereas the last section is dedicated to discussion and outlook.

2. THE PROSPECT

Jacques-Barthélemy Micheli du Crest (1690–1766), the creator of the Prospect, was an outstanding personality with a tragic faith, which has been portrayed a lot in literature and described in great detail in a historiographic novel by the Swiss writer P. Meier [8]. Born in Geneva into a noble family originating from Lucca in Tuscany, the young Micheli was at first a socially approved person. However, he started discussion about Geneva's fortresses and came more and more deeply into trouble with the authorities in his town. As a consequence, he was even sentenced to death in absence. The execution was conducted symbolically on behalf of an image (*in effigie*) because Micheli had fled to France. There, he was in a circle with renowned scientists of the day and kept up with newest developments in physics (thermo- and barometry), astronomy, geodesy and cartography.

Homesick, he returned close to his home region, but not directly to Geneva where it was still dangerous for him to stay. In 1746, on initiative of the authorities of Geneva, Zurich and Bern he was arrested in Neuenburg and brought to Aarburg castle for the first time. After some months in home arrest in Bern, he was sued to life sentence in Aarburg castle. There, he wrote an essay of several pages on June 26, 1754, containing a concept of land survey. This concept comprises

a two-step approach with a land survey on the one side (base measurements, triangulation nets, topographic maps in smaller scales) and detailed surveys with fine-scaled maps [3, pp. 97–101]. Despite his small operating range, Micheli realized the most possible in his situation with primitive instruments: From autumn 1754 to spring 1755, he made a central perspective projection of his field of view, in fact the projection of a portion of the surrounding landscape onto a handy piece of paper, in a well-defined coordinate system of azimuths and height measurements for the determination of the elevations above sea level of 40 alpine summits which limited the distanced field of view. This was the creation of the Prospect. Originally, Micheli entitled it as “Prospect Geometrique des Montagnes neigées, dittes Gletscher, telles qu’on les découvre en tems favorable, depuis le Chateau d’Arbourg, dans les territoires des Grisons, du Canton d’Ury, et de l’Oberland du Canton Berne” (geometric view of the snow-covered mountains, called glaciers, as they are visible in good weather conditions from the Aarburg castle, in the regions of Graubünden, of the canton Uri, and of the oberland in the canton Bern, see Figure 4).

Micheli documented his procedures with a small “Avertissement” (legend) directly on the Prospect, and had a separate “Mémoire” of four printed pages [9]. By means of the latter, the main characteristics of his work may be explained, consisting mainly of the measurements of azimuths and elevations.

Following his description, Micheli constructed the landscape drawing using simple aids in the principle of the measuring table. He drew a quarter circle using a compass and partitioned it “relatively precise”. However, he did not tell anything about this partitioning in detail. Moreover, the quarter circle has not been discovered yet, thus there are only assumptions about this task. Micheli may have partitioned the quarter circle further with the compass, using Archimedes’ pragmatic method.

In order to orientate the quarter circle, Micheli

determined the position of the South direction, the meridian, twice by observing the polar star using a plummet. He calculated the timestamp of its upper and lower meridian transitions using the “*Connaissance des Temps*”, the first astronomical almanach which had been published since 1679. He marked the mean direction in field using a wooden “meridian pole” (French “*Poteau du Méridien*”).

In this way, Micheli determined the South azimuths of the summits in a sector of 44° East (equal to a North azimuth of 136°) and 11° West (equal to an azimuth of 191°), w.r.t. the meridian through Aarburg. The field of view fits therefore well into a quarter circle.

With the azimuths of the Prospect given in regular intervals, the mapping corresponds to a cylindrical projection: In the front view, the landscape is projected onto a vertical cylinder which is developed in the image plane (Figure 1).

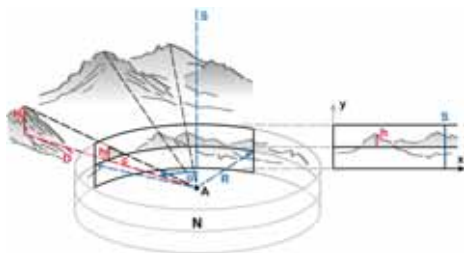


Figure 1: The projection model of the Prospect as a projection of the landscape onto a circular cylinder with vertical axis. Micheli's observed values are the azimuth α w.r.t. the South meridian S and the reading h on the rod w.r.t. the local level N defined by the water level. The elevation H can be determined from the distance D (taken from the map), h and radius c . From [11].

The core element of the whole campaign was the determination of the elevation differences. For this purpose, an extremely simple measuring device was used: A wooden roof gutter of 24 feet (around 7.8m) length, was closed thoroughly on both sides, filled up with water and served

in this way as a giant level, directed in azimuth towards the summits whose elevations were to be determined. The observer placed himself behind this level opposite to the mountains. On the other end of the level, a helper moved a rod of one foot length upwards, perpendicularly to the water surface, until the top of the rod was collinear with the observer's eye and the respective summit. Micheli reports that he usually had the observations done by two very clear-sighted people, each observation was performed several times from far and next to the level. He estimated the measurement error to be less than half a line, corresponding to approximately half an arc second. The “*Hauteurs sur le Niveau*” determined in this way are given in “*Pouces*” and “*Lignes*” in the Prospect.

The elevation differences of the summits could in this way easily be calculated using the rule of proportion: the height of the rod above the water surface of the level gives the opposite leg in an rectangular triangle, the length of the level gives the adjacent leg; together with the distance to the summit, their elevation relative to the observer's position can now be determined.

Micheli knew that the distances to the summits were the main problem in his task. He took the distances out of Scheuchzer's map, published in 1712 in four map sheets as “*Nova Helvetiae tabula geographica*”. He determined its scale by setting the length of one degree in the local latitude in relation to the value published by J. Picard in his book “*La Mesure de la Terre*”. In this way, he calculated the distance from Basel to Geneva and compared it with a newer map from the “*Connaissance des Temps*”. He recognized that he had to reduce the distances in Scheuchzer's map by more than a sixth in order to make them compatible with newest findings. The “*Distances sur la carte de Scheuchzer*” are given in “*Pouces*” and “*Lignes*” above the readings of the vertical rod below the upper scale bar in azimuth (Figure 4). However, Micheli could not eliminate the big interior distortions of Scheuchzer's map with this scale correction.

The influence of Earth curvature corresponding to the respective distance of the summit was determined by means of a table which was based on the principles published in the book “La Mesure de la Terre” by J. Picard in 1671. Micheli did not account for the influence of refraction.

He determined the elevation of his “Niveau” using a long series of observations on a self-constructed barometre. This year-long series resulted in a value of 237 Toises (461.9m). Micheli added the elevation difference to the local elevation of Aarburg and also the amount of Earth curvature. In this way, he determined the elevation above sea level of the summits which are listed in a table in the lower part of the Prospect.

In accordance with his modern way of thinking, Micheli wanted to have his Prospect reproduced and to have his work discussed among prominent scientists. He originally thought about engraving the Prospect in Paris, the copperplate engraving was though made by Tobias Conrad Lotter (1717–1777) in Augsburg.

From top to bottom, the Prospect contains the following main elements:

- Upper scale bar in azimuth
- distance to the respective summit in Scheuchzer’s map (in inches/lines)
- reading on the vertical rod with respect to the local level of Aarburg (in inches/lines)
- horizontal line at the height of 12 inches as upper limit of the landscape image
- landscape image with letters, numbers and symbols for identifying the summits
- local horizon line as lower limit of the landscape image
- lower scale bar following the same principle as the upper one
- four text fields with names of the summits, their elevations and a field with the “Avertissement” (short description).

All elements but the last are visible in Figure 4. The Prospect represents a complete central perspective in a cylindrical coordinate system. Its datum is given by the origin, being the perspective centre, the local meridian, and the local level determining both main axes.

Micheli’s measurements are defined in the French system, based on the “toise de Pérou” determined in a meridian arc measurement in Peru [7, p. 478]. A toise (fathom) corresponds to a length of 1.949m and is subdivided into 6 pieds (feet; 1 pied = 324.84mm), their further subdivided into 12 pouces (inches; 1 pouce = 27.07mm) and each pouce finally subdivided into 12 lignes (lines; 1 ligne = 2.26mm). In this way, 1 toise equals 864 lignes.

3. METHOD OF ANALYSIS

The geometric aspects of the Prospect had been investigated for the first time in 1995 [10]. As digital analysis methods have offered increasing possibilities since then, the analysis was refined in 2011 [11]. It is the basis of the herein presented study, relying in the determination of both image coordinates within a Geographic Information System (GIS) and subsequent evaluation within the photogrammetric bundle block adjustment software ORIENT¹. Thus, Micheli’s work is analysed with the software probably fitting best to the mapping strategy he used at that time. The most important prerequisite for this procedure was the scanning of the Prospect in high resolution in order to retrieve the image coordinates of selected points for evaluation in ORIENT.

The current evaluation has been based on a scan done with 600dpi resolution approximately four years after the first mentioned publication [10]. In comparison to the scan used in [10], distortions of up to 0.2% appeared; they can be explained by paper deformation.

The image coordinates of the following point

¹http://www.ipf.tuwien.ac.at/products/produktinfo/orient/html_hjk/orient.html, last accessed: June 12, 2014

groups were measured in ArcMap²:

- 1 – 39: Points labelled with letters in the Prospect
- 91 – 96: Churches located close to the projection centre, most of them reference points of the Swiss cadastre triangulation
- 101 – 134: further identifiable points in the landscape which were not labelled by Micheli
- 1136 – 1191: lower scale bar in azimuth (geographic North direction = 0° (degrees); the number corresponds to 1000 plus the North azimuth in degrees)
- 2136 – 2191: upper scale bar in azimuth (the number corresponds to 2000 plus the North azimuth in degrees)
- 3001 – 3039: directions to the points 1 – 39
- 4000 – 4012: left scale bar in elevation (the number corresponds to 4000 plus the line index)
- 4100 – 4112: right scale bar in elevation (the number corresponds to 4100 plus the line index)

The coordinates of the points in the first three groups were determined by means of the current official map in scale 1 : 25 000; their standard deviation a priori was set to 5m in x and y coordinates and 1m in elevation. Most points of the second group are reference points of the Swiss cadastre triangulation so that their coordinates are given in high accuracy; their standard deviation a priori was assumed to be $\pm 0.1\text{m}$ in the single coordinates. The heights above ground of these points were further determined photogrammetrically at swisstopo.

A already stated, the numerical analysis was performed in ORIENT. It is a universal software

package for photogrammetric point determination, in which several forms of observations and parameters can be analysed simultaneously [4, 5]:

- Image points from perspective images (x, y) and line scanners $(x = t, y)$,
- 3D model points (x, y, z) ,
- polar points (horizontal direction, zenith angle, distance),
- reference points (X, Y, Z) and
- fictitious observations (planes and lines, surfaces and curves (polynomials), spline curves)

as observations and

- transformation parameters (calibration, orientation),
- additional parameters for image deformation and shape of objects and
- parameters of scanner trajectories (polynomials and splines)

as parameters to be determined or known a priori.

The current special case of a perspective cylindrical projection can be seen as a vertical sensor line rotating around the vertical cylinder axis. The sensor line is a generatrix of the cylinder; the South meridian gives the reference for the generatrices.

The quantitative image analysis using ORIENT is aimed at the reconstruction of the mapping geometry of the prospect. Moreover, the accuracy of this mapping is to be assessed. For this purpose, it is necessary to determine the parameters of the exterior and interior orientation of the Prospect. The first group of the parameters consists of the planar coordinates and the elevation of the draftsman's position, the orientation of the cylinder axis in space and the azimuth of the South direction in the Prospect.

²<http://www.esri.com/software/arcgis>, last accessed: June 12, 2014

The interior orientation describes the position of the projection centre (i.e. the eye of the draftsman) w.r.t. the imaging plane of the sensor line in the image coordinate system. In our case, only the distance of the projection centre to the image plane is of interest. For a narrow viewing angle, this would be referred to as principal distance (cf. the focal length of a camera) [6].

The representation of single elements of the Prospect in ORIENT is as follows:

Geodetic reference frame: The calibrations were performed in the Swiss geodetic reference frame LV03. The azimuths of Micheli in the Prospect and the ones in LV03 differ by the meridian convergence. It amounts to 0.3382° in Aarburg.

Scale bars in azimuth and elevation: As the Prospect does not allow to assign an elevation angle to the points on the upper and lower scale bar (azimuth), the definition of the points 1136 – 1191 and 2126 – 2196 in object space is just given via their image coordinates (with drawing accuracy). This corresponds to an angular measurement with a theodolite in two horizontal planes. The relative position of this two “theodolites” on a vertical line is formulated as a “GESTALT” constraint in ORIENT with an a-priori accuracy of $\pm 1\text{mm}$. Their vertical offset is an unknown parameter to be determined in the adjustment. The scale bars in elevation “Pouces de Nivellement” (inches) are represented as a sequence of Cartesian 3D points, in analogy to the marks on a vertical ($\pm 1\text{mgon}$) measuring rod. The exaggeration in elevation can be determined by setting the distance of these rods to the projection centre to 24 feet (length of the levelling instrument).

Vertical axis of the cylinder: According to the readings on the measuring rod orthogonally to the water surface, the cylinder axis can be set to be vertical up to the reading accuracy

of the rod. Micheli reported the readings to accurate up to $\pm 0.5''$.

Selection of reference points for reconstruction:

As the reference point groups exhibit different accuracy levels, three selection variants were considered. The first was based on 39 points actually measured by Micheli with high accuracy; the second group additionally consisted of 6 points representing churches and castles in close distances, and the third included 34 additional points which were presumably not measured by Micheli but just drawn in the Prospect.

Stochastic model and quality control: The accuracy of the coordinates of the reference points was set to the values given above; the image coordinate accuracy was set to $\pm 0.1\text{mm}$ for both coordinates in case of well-defined points and to $\pm 0.3\text{mm}$ for flat summits orthogonal to the viewing direction.

4. RESULTS

Variance component analysis allows for comparing these a-priori values with the ones retrieved a posteriori. The a-priori accuracies turned out to be too optimistic. As the official maps are of high quality, practically only the accuracies of the image coordinates are left to be altered [6]. Assumed that Micheli could read the azimuths up to $\pm 0.2\text{mm}$ accuracy at his quarter circle, this corresponds to a mean planar accuracy of $\pm 0.3\text{mm}$ in the Prospect. How precise the copperplate engraving fits to the original drawing has not been investigated yet. Assumed this precision amounted to another $\pm 0.3\text{mm}$ the mean planar accuracy would result to $\pm 0.4 - 0.5\text{mm}$.

The influence of single gross errors on the adjustment results was analysed using *data snooping* [2]. In this technique, the most “suspicious” observation is eliminated and the adjustment is repeated without this observation, resulting in a sequence of solutions for the projection centres which can be visualized (Figure 3). For the dif-

ferent reconstruction variants different paths are retrieved which may not correspond to the error ellipses describing the planar accuracy.

The low accuracy of the first selection of reference points may be due to their geometric configuration close to the dangerous circle of resection. Also the freedom of artistic expression, resulting e.g. in oversized drawings of important buildings or slight misplacements of such, may bias the results if not properly analysed in quality control.

On the other hand, knowledge about the topography of the mapped scene ought to be introduced as additional input to data snooping rather than just relying on stochastics.

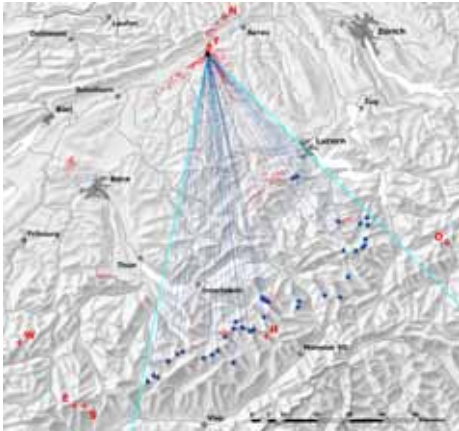


Figure 2: Visibility map containing the summits visible in the Prospect, calculated using SCOP [1] on basis of the digital elevation model DHM25. Visible parts are overlaid in red over the hillshade. From [11].

A simple break of the iteration when a normal distribution of the remaining residuals has been reached may give implausible results—just because the assumed normal distribution need not reflect reality. Let us take the selection variant

²<http://www.swisstopo.admin.ch/internet/swisstopo/en/home/products/height/dhm25.html>, last accessed: June 12, 2014

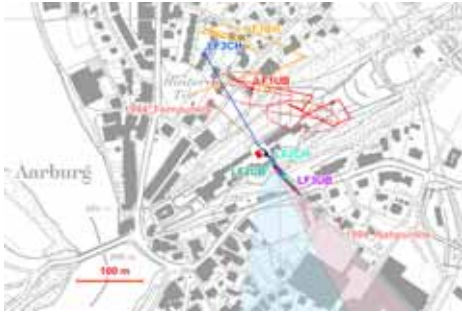


Figure 3: Influence of the selection of reference points on the reconstruction of the projection centre. The two points marked with “1994...” represent the two variants calculated in 1994. Background: Data from the local surveying authority in the canton of Aargau (Übersichtsplan Blatt 94). Numerical values for the error ellipses are given in Table 1. From [11].

“LF3CH” in Figure 3 as an example: elimination of the reference points in the foreground leads to a big jump in the resulting projection centre so that “LF3” corresponds to “LF1” while including more summits. The error of unit weight σ_0 has already been reduced from $\pm 0.59\text{mm}$ to $\pm 0.33\text{mm}$, meeting the expectations. With the jump in the projection centre, the cylinder radius is enlarged by 0.5mm ; however, its standard deviation increases by 20% (from $\pm 0.16\text{mm}$ to $\pm 0.19\text{mm}$). I.e., essential information has been lost.

5. DISCUSSION AND CONCLUSION

The repeated intensive analysis of Micheli’s mapping strategy confirmed the results of the years 1994/95 to a large part. The Prospect, stemming from the age of enlightenment, resolved a main scientific question of that time in an inventive but still effective manner—well adjusted to the difficult situation of its creator— and reached a surprisingly high accuracy.

Data snooping in the “LF1” reference point selection resulted in a standard deviation in the im-

Table 1: Change in size of the semi-axes of error ellipses (A, B [m]) w.r.t. elimination of reference points. Left: before elimination, right: after elimination of the point given in the rightmost column. The σ_0 [mm] is the error of unit weight in the image a posteriori. “Variant” refers to the solutions shown in Figure 3.

variant	A	B	σ_0	A	B	σ_0	elim.
LF1UB	75	30	.284	59	19	.209	11
LF2UB	23	2.6	.256	18	1.8	.229	15
LF3UB	34	3.9	.537	18	2.1	.298	26

age of $\pm 0.2\text{mm}$ after 15 coordinate eliminations out of 39 reference points (originally $\pm 0.3\text{mm}$ at the start). This gives clear evidence that Micheli reached his goal of a “geometrique” representation. In the “LF2” selection variant, this standard deviation decreased from $\pm 0.4\text{mm}$ to $\pm 0.25\text{mm}$, and in “LF3” from $\pm 0.59\text{mm}$ to $\pm 0.4\text{mm}$. These values result from a combination of influences, e.g. the original measurement, the copperplate engraving, random paper distortion, digitization with different scanners.

In addition to the results already given in [10], new findings were made during the enhanced analysis in ORIENT. First, Micheli may have built a model of his measuring device in the scale 1 : 12. As the radius R resulted to $657.4 \pm 0.12\text{mm}$, $657.2 \pm 0.13\text{mm}$ and $656.8 \pm 0.18\text{mm}$ in the selection variants LF1, LF2 and LF3, resp., this corresponds up to 1.1% to 24 inches = 649.7mm. The length of the level was 24 feet. Given that Micheli wanted to map 1 inch of the reading to 1 line in the Prospect (factor 1 : 12), for one calculation variant the scale bar in elevation on the left would have been exaggerated by a factor of 1.073 ± 0.016 and the right one by 1.057 ± 0.016 . The principal distance c varies from 660mm to 710mm in the different variants of selected and eliminated reference points.

Another aspect is the meridian deviation: it is the direction to Micheli’s South direction, corrected by the meridian convergence. Re-projected to the projection plane, values from

-0.3 to -0.7mm in LF2 and -0.6 to -0.9mm in LF3 were retrieved. The common negative sign indicates a slight bias towards the South direction in the mapping frame. A deviation of 0.03° corresponds to 0.03mm in the Prospect. The Student t -factor resulted to values from 0.5 to 1.5; thus, a significant deviation from Micheli’s Polaris measurement could not be proved.

Summarizing the above findings, Micheli’s Prospect Geometrique can be considered as one of the first remarkable geodetic-topographic works, even in its central-perspective conception. It belongs to the initial works fostering the development of the proverbial Swiss precision in this field and of the international reputation of Swiss cartography.

ACKNOWLEDGMENTS

The authors want to express their thanks for the graphical design and image overlays to Michael Pfanner (Figure 1), Nicolai Lanz (Figure 2) and Adrian Bohlen (Figures 5 and 6) from swisstopo.

REFERENCES

- [1] SCOP++ product information. <http://photo.geo.tuwien.ac.at/software/scop/>, 2014. Department of Geodesy and Geoinformation, Research Group Photogrammetry. Last accessed on June 12, 2014.
- [2] W. Förstner. The reliability of block triangulation. In *Proceedings of the 38th Photogrammetric Week*, Schriftenreihe des Instituts für Photogrammetrie der Universität Stuttgart, pages 225–242. Stuttgart, Germany, 1982.
- [3] J. H. Graf. *Das Leben und Wirken des Physikers und Geodäten Jaques Barthélemy Micheli du Crest aus Genf, Staatsgefangener des alten Bern von 1746 bis 1766*. Wyss, Bern, 1890.
- [4] H. Kager. Orient: A universal photogrammetric adjustment system. In A. Grün and

- H. Kahmen, editors, *Optical 3-D Measurement Techniques. Applications in inspection, quality control and robotics*, pages 447–455. Wichmann Verlag, 1989.
- [5] H. Kager. *Orient User Manual*. Institute of Photogrammetry and Remote Sensing, Vienna University of Technology, Austria, 2000.
- [6] K. Kraus. *Photogrammetry – Geometry from Images and Laser Scans*. De Gruyter, Berlin – New York, 2 edition, 2007.
- [7] I. Kretschmer, J. Dörflinger, and F. Wawrik. *Lexikon zur Geschichte der Kartographie. Von den Anfängen bis zum Ersten Weltkrieg*, volume C/1 and C/2. Deuticke, Vienna, Austria, 1986.
- [8] P. Meier. *Die Einsamkeit des Staatsgefangenen Micheli du Crest. Eine Geschichte von Freiheit, Physik und Demokratie*. Pendo, Zurich – Munich, 1999.
- [9] J.-B. Micheli du Crest. *Mémoire pour l'explication du Prospect des Montagnes neigeées que l'on voit du château d'Aarbourg*. 4 pages.
- [10] M. Rickenbacher. Das Alpenpanorama von Micheli du Crest – Frucht eines Versuches zur Vermessung der Schweiz im Jahre 1754. *Cartographica Helvetica*, (Sonderheft 8): 24, 1995. With facsimile of the Prospect Geometrique and reconstruction of the panorama by means of the DHM 25.
- [11] M. Rickenbacher and H. Kager. Geometrische Bildanalysen am Beispiel des „Prospect Geometrique“ von Micheli du Crest von 1754. In *Festschrift für Univ.-Prof. Dr.-Ing. Kurt Brunner anlässlich des Ausscheidens aus dem aktiven Dienst*, number 87 in Schriftenreihe des Instituts für Geodäsie der Universität der Bundeswehr München, pages 197–212. 2012.

ABOUT THE AUTHORS

1. Helmut Kager: born in 1950. Studies in geodesy at Vienna University of Technology. Dipl.-Ing. (MSc) 1974, Dr.techn. (PhD) 1980. Since 1974 with the Institute of Photogrammetry and Remote Sensing (since 2012 within the Department of Geodesy and Geoinformation) at Vienna University of Technology. Contact: helmut.kager@geo.tuwien.ac.at.
2. Martin Rickenbacher: born in 1954. Studies in geodesy at ETH Zurich and history at Basel University; Dipl.-Ing. ETH. Dr.phil. (PhD) in history 2009. Scientist at the Swiss Federal Office of Topography *swisstopo* and leader of the Working Group on the History of Cartography in the Swiss Society of Cartography (SSC). Eratosthenes Award 2011. Contact: martin.rickenbacher@swisstopo.ch.
3. Andreas Roncat: born in 1981, studies in geodesy, geometry and computer science at Vienna University of Technology and University of Vienna. Dipl.-Ing. (MSc) 2006. Currently Research Associate and PhD Student at the Research Groups Photogrammetry and Remote Sensing, Department of Geodesy and Geoinformation, Vienna University of Technology. Contact: andreas.roncat@geo.tuwien.ac.at.

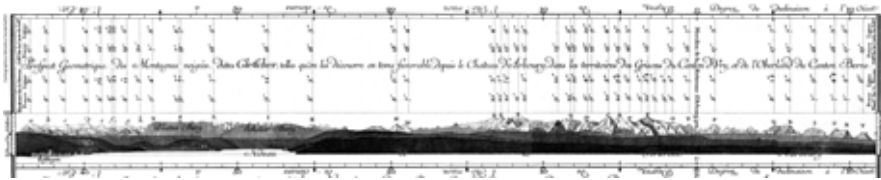


Figure 4: Upper part of the Prospect (facsimile).

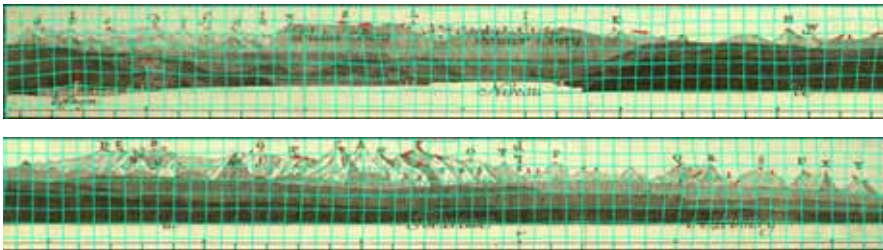


Figure 5: Visualization of the deformation analysis. Grid width: 30' for vertical lines and 2 pouces (inches) for horizontal lines. Residual vectors are given in scale 1 : 1. Top: left part, bottom: right part of the Prospect. From [11].

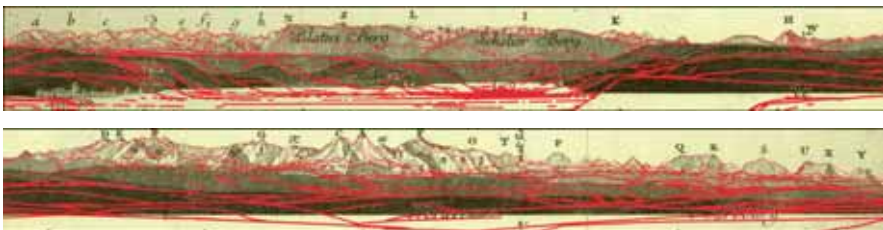


Figure 6: Graphical comparison by overlaying the Prospect with silhouettes generated with SCOP [1] and the digital elevation model DHM25. The vertical exaggeration of the Prospect was considered. Top: left part, bottom: right part of the Prospect. From [11].

THE GRAPHIC INTERPRETATION OF THE CO-OCCURRENCE NETWORK DIAGRAM IN DIFFERENT WRITING STYLES

Miya YAMADE and Hirokazu ABE

Osaka University, Japan

ABSTRACT: The co-occurrence network diagram is one of the methods of Text Mining, which has been used in language analysis and marketing research. However, at present, interpretation of the co-occurrence network diagram has not been standardised. This study aims to explore a new interpretive procedure for the co-occurrence network diagram through a graphical approach using seven different types of writing: novels, editorial articles, newspaper articles, theses, essays, descriptions, and blog articles. Results are as follows: 1) Theses and descriptions had higher average values for the evaluation items than other types of writing. Therefore, there are many co-occurrence words that are involved in the sentences in these types of writing. 2) Novels and essays had many sentences with words used only once per sentence. 3) Blog articles were similar to novels and essays. However, the number of communities showed that blog articles were also similar to newspaper articles suggesting that blog articles have few topic groups in a paragraph. 4) The evaluation items showed some similarities between editorial articles and newspaper articles. Moreover, editorial articles and newspaper articles had many connections with three words in a sentence.

Keywords: Applied geometry, co-occurrence network diagram, text mining, type of writing, fixed figure

1. INTRODUCTION

Text mining is a quantitative text analysis technique that has been used in language analysis and marketing research. In recent years, it has been used to analyse free description-type questionnaire responses, but its use is expected to expand to various research applications in the future. One of the methods of analysis in text mining is the co-occurrence network diagram. The co-occurrence network diagram visualises the connections, patterns and similarities between the words that appear in a sentence and has often been used for writing analysis. In a previous study, Matsumoto et al. [1] investigated drawing a co-occurrence network diagram using text data obtained from office environmental research to improve the office environment. Another study [2]

considered the evolution of an argument by observing the spread of ideas in a time series and examining the process of discussion from the increase in words contained in the co-occurrence network diagram. A study by Okada et al. Nagano et al. [3] interpreted the details of a document from the title using a divided co-occurrence network diagram that focused on the occurrence of words connected in the same community. Fukui et al. [4] compared the variety and frequency of words occurring in a sentence to interpret the details of text obtained from an interview survey. These previous studies interpreted the contents of the original text from clue words that appeared in the co-occurrence network diagram or used the co-occurrence network diagram to identify trends of whole sentences. However, a standardised method of interpreting the

co-occurrence network diagram has not yet been established. This paper aims to explore a new procedure for interpreting the co-occurrence network diagram through a graphical approach by analysing seven different types of writing.

2. METHODOLOGY

2.1 RESEARCH METHOD

Figure 1 shows an example of the co-occurrence network diagram indicating co-occurrence of words and similar patterns of appearance. Circles represent the words and lines show the connections between words. Strong co-occurrences are indicated by thick lines and the most frequent occurrences are represented by a large circle.

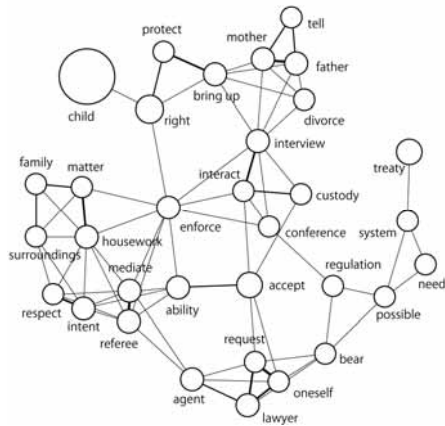


Figure 1: Co-occurrence network diagram for editorial article No. 5.

One method of interpreting such diagrams is to divide words into communities (groups with a common attribute) with nodes and links based on the relationships between phrases. In this study, we used the KH Coder [5] to draw co-occurrence network diagrams and to identify the differences between the diagrams for the

following seven types of writing: novels, editorial articles, newspaper articles, theses, essays, descriptions, and blog articles. Data were collected from the Internet and publications with around 180 (range 107-219) words. Novels (n=9) were extracted from the digital library, Aozora Bunko. Editorial articles (n=11) and newspaper articles (n=2) were extracted from the Asahi, Tokyo, Yomiuri, Nikkei, and Mainichi newspapers. Theses (n=6) were taken from AIJ (Architectural Institute of Japan), essays (n=5) from women's magazines and descriptions (n=2) from home appliance instructions. Blog articles (n=5) were collected from a Google blog search using one theme. We implemented subgraph detection by modularity by dividing the co-occurrence network diagrams into several communities and analysed the characteristics of the seven types of writing from the numbers of nodes, links, communities, and fixed figures as well as the area of the fixed figures.

2.2 CREATING A CO-OCCURRENCE NETWORK DIAGRAM

Co-occurrence network diagrams were drawn using the nine parts of speech "nouns", "verbs" and "adjectives" from sentences in which the same word appears more than once. In addition, draw nodes that lead to links when a word has less than 60 links. On the other hand, draw nodes that only follow minimum Jaccard index when a word has over 60 links. Jaccard index is a coefficient using for comparing similarity and strength of connection of the words. Keyword $n(X)$ is the number of hits of keyword X , Jaccard index between A and B is to:

$$Jaccard(A, B) = \frac{n(A \cap B)}{n(A \cup B)}$$

The subgraph detection used in this paper was based on Clauset's [6] and Newman's [7] theory of modularity. Modularity is a helpful indicator to extract communities from networks.

Clauset and Newman define modularity as:

$$Q = \sum_i (e_{ii} - a_i^2)$$

3. RESULTS AND DISCUSSION

3.1 CO-OCCURRENCE NETWORK DIAGRAM

Figures 2 show the co-occurrence network

diagrams of the seven types of writing that resulted from subgraph detection by modularity. Theses and descriptions were found to have a lot of co-occurring words and phrases in a sentence (i.e., many nodes and links), and they are more complex and intertwined than other types of writing. However, novels, blogs, and essays showed few nodes and links. Moreover, some differences were seen in the number of

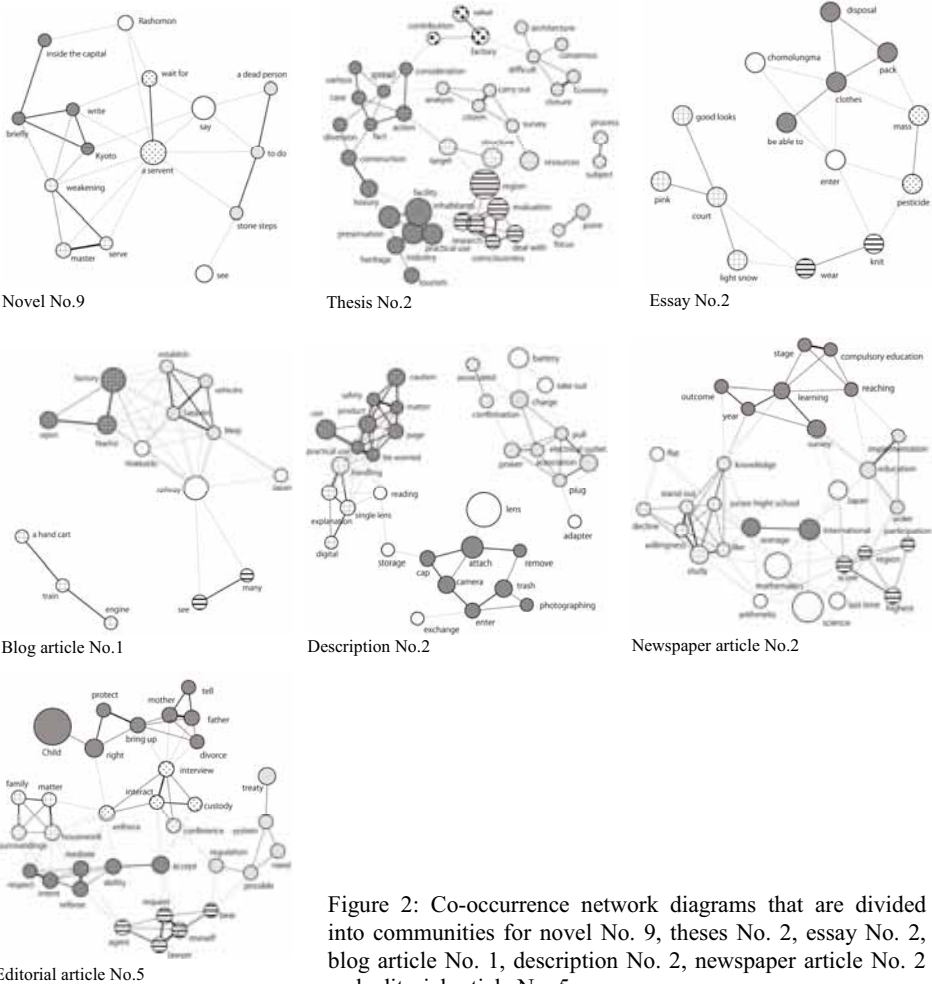


Figure 2: Co-occurrence network diagrams that are divided into communities for novel No. 9, theses No. 2, essay No. 2, blog article No. 1, description No. 2, newspaper article No. 2 and editorial article No. 5.

nodes in descriptions and blog articles, but the nodes were intertwined in a complex with many links as part of a community. Newspaper and editorial articles were complexly intertwined; however, there was no meaning in the placement of each node. The thickness of the lines in co-occurrence network diagrams only shows the strength of words. Therefore, in our interpretation method, we try to represent the strength and weakness of words by the length of the links.

Co-occurrence network diagrams are formed as aggregate data of nodes. Figure 3A shows a triangle called a fixed figure with three nodes and three links that are fixed. On the other hand, Figure 3B and 3C have four nodes and four links and are not fixed.

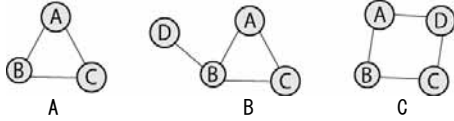


Figure 3: Diagrams of a fixed figure (A) and figures with four links and four nodes (B and C) that are not fixed

This study focused on fixed figures created by the nodes and the links in a community to interpret the co-occurrence network diagram graphically. Co-occurrence network diagrams represent the strength of the connections of the three words in a fixed figure from the characteristics of the co-occurrence network diagram. Therefore, the size of a fixed figure determines the strength of the connections of the three words.

3.2 GRAPHICAL INTERPRETATION OF CO-OCCURRENCE NETWORK

We analysed the characteristics of each writing type using the number of words, the total number of nodes, the total number of links, the number of drawing nodes, the number of drawing links, the number of communities, the

number of fixed figures, and the area of fixed figures as evaluation items (Table 1) and then the values were averaged (Table 2). In order to reduce the influence of differences between individual words, each value was divided by the number of words.

Table 1: Evaluation of the co-occurrence network

Type of writing	Number of words	Total number of nodes	Total number of links	Number of drawing nodes	Number of drawing links	Number of communities	Number of fixed figures	Area of fixed figures
Novel No.1	189	37 (0.20)	27 (0.14)	15 (0.18)	27 (0.14)	4	5 (0.03)	0.3 (0.002)
Novel No.2	178	21 (0.12)	21 (0.12)	11 (0.06)	21 (0.12)	4	1 (0.01)	0.01 (0.000)
Novel No.3	184	20 (0.11)	33 (0.18)	13 (0.07)	33 (0.18)	3	7 (0.04)	0.18 (0.001)
Novel No.4	194	17 (0.09)	20 (0.10)	10 (0.05)	20 (0.10)	3	2 (0.01)	0 (0)
Novel No.5	197	19 (0.10)	7 (0.04)	6 (0.03)	7 (0.04)	3	0 (0)	0 (0)
Novel No.6	219	31 (0.14)	34 (0.16)	19 (0.08)	34 (0.16)	4	10 (0.05)	0.37 (0.002)
Novel No.7	218	28 (0.13)	78 (0.35)	27 (0.12)	81 (0.28)	6	25 (0.11)	0.83 (0.004)
Novel No.8	202	18 (0.09)	46 (0.23)	21 (0.10)	46 (0.23)	5	11 (0.05)	0.26 (0.001)
Novel No.9	205	26 (0.13)	29 (0.14)	15 (0.07)	29 (0.14)	4	2 (0.01)	0.05 (0)
Editorial article No.1	152	24 (0.16)	71 (0.47)	23 (0.15)	68 (0.45)	5	14 (0.09)	0.43 (0.003)
Editorial article No.2	165	21 (0.13)	65 (0.39)	21 (0.13)	65 (0.39)	4	16 (0.10)	0.32 (0.002)
Editorial article No.3	188	36 (0.19)	161 (0.86)	36 (0.19)	90 (0.48)	6	37 (0.20)	0.8 (0.004)
Editorial article No.4	141	34 (0.24)	198 (1.40)	31 (0.22)	64 (0.45)	5	27 (0.19)	1.3 (0.009)
Editorial article No.5	171	34 (0.20)	116 (0.68)	33 (0.19)	75 (0.44)	6	27 (0.16)	0.85 (0.005)
Editorial article No.6	171	41 (0.24)	261 (1.53)	40 (0.23)	109 (0.64)	6	55 (0.32)	2.87 (0.017)
Editorial article No.7	173	27 (0.16)	160 (0.92)	27 (0.16)	63 (0.36)	5	15 (0.09)	0.75 (0.004)
Editorial article No.8	164	18 (0.11)	42 (0.26)	17 (0.10)	42 (0.26)	3	11 (0.07)	0.32 (0.002)
Editorial article No.9	150	32 (0.21)	206 (1.37)	32 (0.21)	76 (0.51)	4	37 (0.25)	1.67 (0.011)
Editorial article No.10	174	20 (0.11)	61 (0.35)	20 (0.11)	61 (0.35)	5	10 (0.06)	0.34 (0.002)
Editorial article No.11	201	52 (0.26)	234 (1.16)	47 (0.23)	75 (0.37)	9	32 (0.16)	0.79 (0.004)
Thesis No.1	195	60 (0.31)	835 (4.28)	47 (0.24)	82 (0.42)	8	47 (0.24)	14.47 (0.074)
Thesis No.2	163	72 (0.44)	561 (3.44)	40 (0.25)	68 (0.42)	8	21 (0.13)	2.93 (0.018)
Thesis No.3	179	50 (0.44)	508 (2.84)	48 (0.27)	107 (0.60)	6	50 (0.28)	6.17 (0.034)
Thesis No.4	113	25 (0.22)	208 (1.85)	25 (0.22)	80 (0.71)	4	22 (0.19)	0.55 (0.005)
Thesis No.5	107	18 (0.17)	102 (0.95)	18 (0.17)	69 (0.64)	3	19 (0.18)	1.06 (0.010)
Thesis No.6	162	37 (0.23)	354 (2.19)	36 (0.22)	73 (0.45)	3	18 (0.11)	0.97 (0.006)
Essay No.1	185	29 (0.16)	57 (0.31)	29 (0.16)	57 (0.31)	5	17 (0.09)	0.31 (0.002)
Essay No.2	208	24 (0.12)	20 (0.10)	14 (0.07)	20 (0.10)	4	1 (0.005)	0.03 (0)
Essay No.3	129	22 (0.17)	52 (0.41)	22 (0.17)	52 (0.41)	4	13 (0.10)	0.33 (0.003)
Essay No.4	156	33 (0.21)	27 (0.17)	27 (0.17)	27 (0.17)	7	6 (0.04)	0.07 (0)
Essay No.5	84	14 (0.17)	13 (0.15)	12 (0.14)	13 (0.15)	4	1 (0.01)	0.01 (0)
Description No.1	200	56 (0.28)	287 (1.44)	44 (0.22)	76 (0.38)	9	38 (0.19)	5.21 (0.026)
Description No.2	187	36 (0.19)	131 (0.70)	34 (0.18)	72 (0.39)	5	25 (0.13)	1.44 (0.008)
Newspaper No.1	123	17 (0.14)	66 (0.54)	17 (0.14)	60 (0.49)	2	22 (0.18)	1.52 (0.012)
Newspaper No.2	125	29 (0.23)	141 (1.13)	29 (0.23)	79 (0.63)	5	24 (0.19)	0.72 (0.006)
Blog article No.1	138	19 (0.14)	38 (0.28)	15 (0.11)	38 (0.28)	4	5 (0.04)	0.35 (0.003)
Blog article No.2	149	24 (0.16)	59 (0.4)	24 (0.16)	59 (0.4)	4	28 (0.19)	0.85 (0.004)
Blog article No.3	164	24 (0.15)	75 (0.46)	24 (0.15)	64 (0.39)	4	26 (0.16)	0.73 (0.004)
Blog article No.4	162	12 (0.07)	8 (0.05)	8 (0.05)	8 (0.05)	3	1 (0.01)	0 (0)
Blog article No.5	159	16 (0.10)	21 (0.13)	14 (0.08)	21 (0.13)	4	3 (0.02)	0.05 (0)

Numbers in brackets are values divided by the number of words

Table 2: Average values of the co-occurrence network evaluation items

Type of writing	Number of words	Total number of nodes	Total number of links	Number of drawing nodes	Number of drawing links	Number of communities	Number of fixed figures	Area of fixed figures
Novel	198.44	24.1 (0.12)	32.6 (0.15)	15.1 (0.08)	30.9 (0.16)	4.0	7.0 (0.04)	0.22 (0.001)
Editorial article	168.16	30.8 (0.18)	143.2 (0.85)	29.7 (0.18)	71.7 (0.43)	5.3	25.5 (0.15)	0.05 (0.008)
Thesis	153.07	43.7 (0.29)	428.2 (2.90)	35.7 (0.23)	79.8 (0.52)	5.3	29.5 (0.19)	4.30 (0.029)
Essay	152.2	24.4 (0.16)	33.8 (0.22)	20.8 (0.14)	33.8 (0.22)	4.8	7.8 (0.05)	0.19 (0.001)
Description	193.5	46 (0.24)	209 (1.08)	39 (0.20)	74 (0.38)	7.0	31.5 (0.18)	3.33 (0.017)
Newspaper article	124	23 (0.19)	103.5 (0.83)	23 (0.19)	66.5 (0.56)	3.5	23.0 (0.19)	1.12 (0.009)
Blog article	183.31	19 (0.12)	40.2 (0.26)	17 (0.11)	38 (0.25)	3.8	26.2 (0.17)	0.35 (0.002)

Numbers in brackets are values divided by the number of words

■ Max □ Min

3.2.1 NUMBER OF NODES AND LINKS

Figure 4 shows the percentage of drawing nodes and the number of drawing links. Theses and descriptions had the highest percentage of total drawing nodes and drawing links (≥ 0.24), and newspaper articles, editorial articles, essays, novels, and blog articles had the lowest (0.12-0.16). The average of number of drawing nodes showed a similar trend as the total number of nodes. The percentage of total links was especially high for theses (2.8), while that of descriptions, newspaper articles and editorial articles was almost the same (0.83-0.85). On the other hand, Figure 4 suggests that theses, newspaper articles and editorial articles have a high percentage of drawing links while the percentage of drawing links in novels, essays and blog articles is low (0.3).

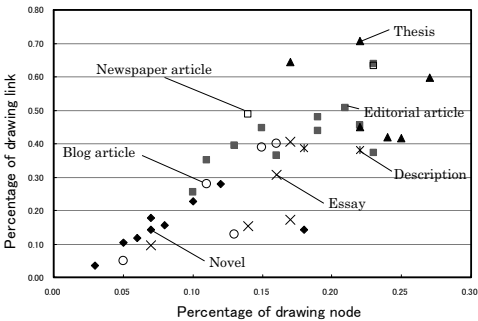


Figure 4: Percentage of drawing nodes and drawing links.

These results suggest that theses and descriptions have many nodes and more co-occurrence words than other types of writing. Moreover, theses, newspaper articles and editorial articles have many drawing links that intimate co-occurrence words and are the most complex. Meanwhile, novels and essays have fewer nodes and links than other types of writing suggesting that the same word does not occur repeatedly in these types of writing.

3.2.2 NUMBER OF COMMUNITIES

Figure 5 shows the number of communities and drawing nodes and Table 2 shows the average number of communities in each type of writing. The highest average number of communities was found in descriptions (7.0). The average number of communities in theses and editorial articles was 5.3, but that of newspaper articles and blog articles was ≤ 3.8 . Thus, the topics discussed in theses are more unified than in those in newspaper articles and blog articles.

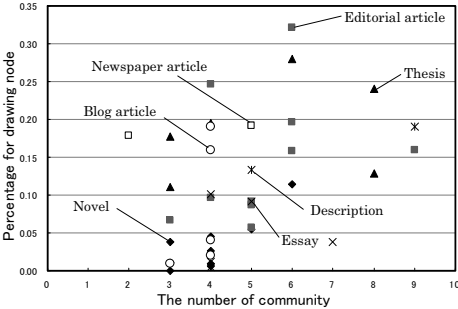


Figure 5: The number of communities and percentage of drawing nodes.

3.2.3 NUMBER OF FIXED FIGURES AND THE AREA OF FIXED FIGURES

Figure 6 shows the percentage of the number of fixed figures and the area of fixed figures. The highest average percentage of the number of fixed figures was for theses and newspaper articles (0.19) while those of blog articles, descriptions and editorial articles were

0.15-0.17. On the other hand, novels and essays had percentages of ≤ 0.05 . The highest average area of fixed figures was for theses (0.028). Descriptions had an average area of 0.017, while blog articles, essays and novels had an average area of 0.002 or less.

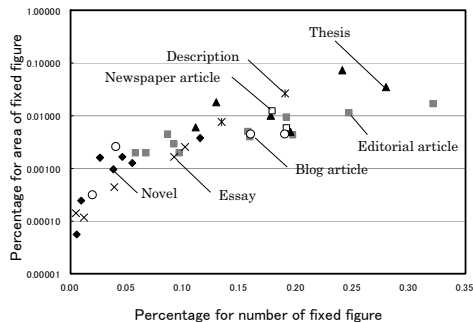


Figure 6: Percentage of number of fixed figures and area of fixed figures.

Thus, theses and newspaper articles had many connections among three words in one sentence. However, theses and descriptions have sentences that are more complex and intertwined than blog articles.

4. CONCLUSIONS

We investigated a new interpretive procedure for co-occurrence network diagrams by examining seven different and distinct types of writing. Results are as follows:

1. Theses and descriptions had higher average values for the evaluation items than other types of writing. Therefore, there are many co-occurrence words that are involved in the sentences in these types of writing.
2. Novels and essays had many sentences with words used only once per sentence. Blog articles were similar to novels and essays. However, the number of communities showed that blog articles were also similar to newspaper articles

suggesting that blog articles have few topic groups in a paragraph.

3. The evaluation items showed some similarities between editorial articles and newspaper articles. Moreover, editorial articles and newspaper articles had many connections with three words in a sentence.

Although it might be necessary to compare other types of writing, the current study offers a new procedure for interpreting the co-occurrence network diagram.

REFERENCES

- [1] Matsumoto N, Honda H, Matsumoto Y, Kidosaki N, and Naka R: Requirements of Office Environment from Viewpoints of Workers (Part2): Analysis of Evaluation Tendencies, Using Text-Mining Approach, Architectural Institute of Japan (Hokuriku), pages 411–412. 2010.9.
- [2] Okada Y, Taniguchi M, Matsumoto Y, Kayahara T, Jinushi H, and Naka. R: A Study on New Communication Media for Meeting Environment: Evaluation of Utterance Analysis and Co-occurrence Network Analysis, Architectural Institute of Japan (Hokuriku), pages 535–536. 2009.8.
- [3] Eino S, Oneyama Y, and Oguchi T: Study on Analysis Method of Free Answers in Questionnaire Survey Using Morphological Analysis: Case Study of Attitude Survey for Utility of Tram, Journal of JSCE, pages 973–981. , 68(5), 2012.
- [4] Fukui M, Abe H, and Hashitera T: A Primary Contractor's Role in Conservation of Industrial Heritage: Cases of textile mill in Osaka and Hyogo, Journal of Architecture and Planning (2013), Vol.78, No. 687, pages 1067–1076.
- [5] Higuti K: KH Coder 2.x Tutorial. KH Coder (2011), pages 60–61.
- [6] Clauset, A., M. E. J. Newman & C. Moor: Finding community structure in very large

networks, Physical Review E, 70(6): 066111(2004), pages 066111-1–066111-6

- [7] M. E. J. Newman: Fast algorithm for detecting community structure in networks, Physical Review E, 69: 066133(2004), pages 066133-1–066133-5.

ABOUT THE AUTHORS

1. Miya YAMADE, Dr. Eng., is a Research Fellow at Graduate School of Engineering, Osaka University, Yamadaoka 2-1, Suitashi, Osaka, 565-0871, JAPAN

2. Hirokazu ABE, Dr. Eng., is a Professor at Cybermedia Center, Osaka University, Mihogaoka, Ibarakishi, Osaka, 567-0047, JAPAN

HEXAHEDRAL MESH GENERATION USING ADAPTIVE SAMPLING BASED ON SHARP FEATURES

Yusuke IMAI¹, Hiroyuki HIRAOKA¹, and Hiroshi KAWAHARADA²

¹ Chuo University, Japan ² Yokohama National University, Japan

ABSTRACT: Many manufacturers now use computer-aided design (CAD) for CNC machining, simulations, press working, etc. They use CAD models for simulations because the cost of performance simulations is lower than that of tests using actual products. In this paper, we consider hexahedral meshes for finite element analysis, because simulations with such meshes are more accurate than those with tetrahedral meshes. Our aim is to automatically generate hexahedral meshes with sharp features that precisely represent the corresponding features of a target shape. Our hexahedral mesh generation algorithm is voxel-based. Voxels (initial hexahedral mesh) in our algorithm is generated by Polymender. Polymender generates voxels whose normal vectors are parallel to x, y, z axes. Thus, target surface's attitude should be along to x, y, z axes. Therefore, we apply rotation algorithm for target surface. After that, we used up-sampling to increase the number of points on the target surface mesh in order to increase the number of candidates for the corresponding point. To represent sharp features, we increase candidates near the sharp features. In this paper, we propose rotation algorithm and new up-sampling method for target surface. Furthermore, we present some experimental results using our method.

Keywords: Hexahedral mesh, Up-sampling, Sharp feature, Adaptive, Voxel.

1. INTRODUCTION

In manufacturing, the cost of PC simulations is lower than that of testing actual prototypes. Thus, most manufacturers run PC simulations that require volume meshes. Until around a decade ago, the simulation process started from surface meshes made using computer-aided design (CAD) software. However, there were differences between the shape of the actual products and the CAD model by a result of manufacturing factors such as springback, diecasting, etc. Even if manufacturers used metal dies at actual diecasting whose shapes are identical to the CAD models, the parts obtained from certain processes did not have the same shape as the CAD model. This difference makes various simulations using CAD models unrealistic. Today, the simulation process starts from point clouds scanned from the actual products. This process is called reverse engineering.

The tetrahedral/hexahedral meshes generated from such point clouds directly affect the results given by finite element method (FEM) analysis. Hexahedral meshes are superior to tetrahedral meshes (see Figure 1) in terms of accurate analysis. Thus, in this paper, we consider hexahedral volume meshes (see Figure 1) whose elements are hexahedral cells only (we call these all-hexahedral meshes), and examine their surface.

In the structural analysis of two objects, the stress peaks near or around the contact regions. Such regions are often sharp features, typically the cusp (such as an edge or point) of an object. Thus, to obtain accurate simulation results, the surface mesh must represent sharp features.

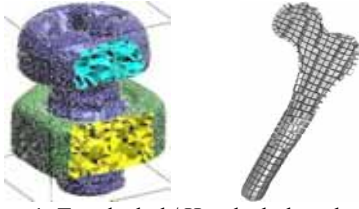


Figure 1: Tetrahedral / Hexahedral meshes.

In this paper, we consider a voxel-based hexahedral mesh generation algorithm [1-2]. In addition, the volume mesh for the FEM must satisfy the constraint that all Jacobians are positive. The Jacobian is a triple scalar product $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$, where \mathbf{a}, \mathbf{b} , and \mathbf{c} are vectors (edges) adjacent to a corner vertex of a cell.

Our ultimate goal is automatic hexahedral mesh generation without negative Jacobians. In this study, we discuss the quadrilateral surface of a hexahedral mesh (The algorithm we investigated in previous studies is also voxel-based [3-4]). Thus, our inputs are the target surface mesh and the quadrilateral mesh that is the surface of the voxel mesh.

Hexahedral mesh generation algorithms can be fully or semi-automatic, but there is no scheme that guarantees all Jacobians will be positive. On the other hand, there is a scheme for tetrahedral meshes that guarantees positive Jacobians. Before considering how to achieve positive Jacobians, we generate quadrilateral surfaces of the all-hexahedral meshes that represent the target surfaces. In this paper, we propose an automatic fitting algorithm with sharp features.

2. OUR PREVIOUS HEXAHEDRAL MESHING ALGORITHM

First, we summarize the previous hexahedral mesh generation algorithm [3-4]. The underlying algorithm can be broken down as follows.

1. Input target surface mesh (see Figure 2).
2. Generate voxels to wrap around the target surface using Polymender [5] (see Figure 3).

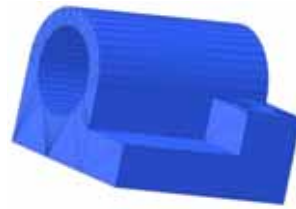


Figure 2: Target surface.

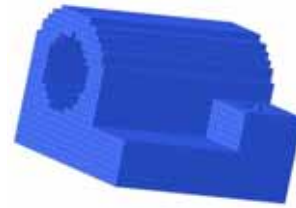


Figure 3: Voxels.



Figure 4: Previous result.

3. Extract the boundary surface of voxels.
4. Fit the boundary surface of voxels (fitted surface is shown in Figure 4[6]).
5. Determine the positions of the inner vertices.
6. Apply post-processing.
7. Output hexahedral mesh.

In step 4 above, we fit the boundary surface of the voxels to the target surface mesh. We call this the fitting part.

3. FITTING ALGORITHM

In this section, we describe our proposed fitting algorithm. Let $\mathbf{v}_i^0 (i = 1, 2, \dots)$ be the positions of the voxel's boundary vertices at step 0 (initial positions). The fitting algorithm is iterative ($k =$

0 to T) [3-4].

Algorithm 1:

1. Determine \mathbf{v}_i^{k+1} ($i=1,2,\dots$), by minimizing the following quadratic energy function, as outlined [7] where \mathbf{v}_i^k denotes the position of boundary vertex \mathbf{v}_i at step k , and \mathbf{w}_i^{k+1} denotes the corresponding point on the input surface to \mathbf{v}_i^k .
2. Apply surface Laplacian smoothing to the $(k+1)$ -th surface. If $k \neq T$, then increase k by 1 and return to step 1.

Let i^* be the set of integer if $j \in i^*$, then \mathbf{v}_j is adjacent to \mathbf{v}_i ((i,j) is an edge of the boundary quadrilateral surface). Let \mathbf{w}_i^{k+1} be the corresponding point on the input surface to \mathbf{v}_i^k . \mathbf{nw}_{it}^k ($t=1,2,\dots,A_i^k$) and \mathbf{nw}_{it}^{k+1} ($t=1,2,\dots,C_i^{k+1}$) be multi-normal vector at \mathbf{v}_i^k on the k -th boundary surface and \mathbf{w}_i^{k+1} on the target surface respectively. Let A_i^k, C_i^{k+1} be the numbers of normal vectors of $\mathbf{nw}_{it}^k, \mathbf{nw}_{it}^{k+1}$, and $M_i^k = A_i^k + C_i^{k+1}$ [6]. Let N_v be the number of boundary vertices. $\|\cdot\|$ denotes the norm of a vector. $\mathbf{a} \cdot \mathbf{b}$ denotes the inner product between vectors \mathbf{a} and \mathbf{b} . T is the user-defined maximum iteration number ($T=30$ in this paper). The minimization of the quadratic energy function can be written as follows [6]:

$$\min_{\mathbf{v}_i^{k+1}} \left[\sum_{j \in i^*} \left\{ c_1 \left\| \left(\mathbf{v}_i^{k+1} - \frac{1}{|i^*|} \sum_{j \in i^*} \mathbf{v}_j^{k+1} \right) - \left(\mathbf{v}_i^k - \frac{1}{|i^*|} \sum_{j \in i^*} \mathbf{v}_j^k \right) \right\|^2 \right\} + c_2 \left\| \mathbf{v}_i^{k+1} - \mathbf{v}_i^k \right\|^2 + c_3 \left\| \mathbf{v}_i^{k+1} - \mathbf{w}_i^{k+1} \right\|^2 \right] + c_4 \left\{ \sum_{(h,l)} \left(\frac{-n\mathbf{w}_{it}^{k+1}}{d_{it}^{k+1}} \cdot n\mathbf{w}_{it}^{k+1} \right) / M_i^k \right\} \quad (1)$$

Here, $c_1=1, c_2=0.5, c_3=2, c_4=0.01$.

After the fitting, we carry out Laplacian smoothing to improve the quality of the resulting all-hexahedral mesh.

4. INITIAL MESH OPERATION

4.1 Initial mesh rotation

Our hexahedral mesh is generated by a voxel-based algorithm. To generate a high quality hexahedral mesh, in this paper, the normal vectors of both the initial voxels and the target surface should lie approximately in the same direction. Thus, we rotate the target surface to maximize the area containing similar normal vectors before the voxel generation.

Initial voxels (Figure 5) are generated by Polymender using target surface that is 45 degree rotation around the z-axis of the target surface (Figure 2). We fit boundary surface of this voxels to target surface, however we could not get high quality mesh (Figure 6). As above, actual input target surface's attitude is various. Thus, we must rotate target surface to get target surface whose normal vectors' directions are same as x, y, z axes.

We set $n\mathbf{f}_p$ ($p=1,2,\dots,N_f$) as facet normal vector of input target surface. Let N_f be the number of target surfaces face. $n\mathbf{f}_p$ is not unit vector. Its norm is the area of the face. Let $\mathbf{n}_{p,axis_i}$ be one of the six direction (unit) vectors of axis $(\pm x, \pm y \pm z)$. Here, we define the evaluation value between $n\mathbf{f}_p$ and $\mathbf{n}_{p,axis_i}$ as follows:

$$\sum_p \left(\max_{axis_i} (n\mathbf{f}_p \cdot \mathbf{n}_{p,axis_i}) \right). \quad (2)$$

In order to do this, we calculate all the inner products between the facet normal vector $n\mathbf{f}_p$ and the direction vector $\mathbf{n}_{p,axis_i}$ that is one of $(\pm x, \pm y \pm z)$ unit vectors (See Table 1). We obtain the maximum of each row. We define evaluation value $\sum_p \left(\max_{axis_i} (n\mathbf{f}_p \cdot \mathbf{n}_{p,axis_i}) \right)$ for rotation of the target surface as the sum of the maximums.

Table 1: Matching table.

	x	$-x$	\dots	z	$-z$	MAX
\mathbf{nf}_1^k	3.1	-3.1		0.2	-0.2	3.1
\mathbf{nf}_2^k	1.8	-1.8		5.7	-5.7	5.7
\vdots						
$\mathbf{nf}_{N_f}^k$	2.0	-2.0		-5.1	5.1	6.4
SUM						145.7

In Table 1, we show the evaluation value as bottom right cell. Let \mathbf{R} be a rotation matrix. Then, $\mathbf{R}(\mathbf{nf}_p)$ ($p=1,2,\dots,N_f$) are rotated facet normal vectors.

Here, we want initial rotation \mathbf{R} that maximize Equation (3).

$$\max_{\mathbf{R}} \left[\sum_p^{N_f} \left(\max_{axis_i} \left(\mathbf{R}(\mathbf{nf}_p) \cdot \mathbf{n}_{p,axis_i} \right) \right) \right] \quad (3)$$

The matrix \mathbf{R} can be expressed by a unit quaternion $\mathbf{q} \equiv (q_0, q_1, q_2, q_3)$, ($q_0^2 + q_1^2 + q_2^2 + q_3^2 = 1$).

$$\mathbf{R} = \begin{pmatrix} q_0^2 + q_1^2 - q_2^2 - q_3^2 & 2(q_1q_2 - q_0q_3) & 2(q_1q_3 + q_0q_2) \\ 2(q_1q_2 + q_0q_3) & q_0^2 - q_1^2 + q_2^2 - q_3^2 & 2(q_2q_3 - q_0q_1) \\ 2(q_1q_3 - q_0q_2) & 2(q_2q_3 + q_0q_1) & q_0^2 - q_1^2 - q_2^2 + q_3^2 \end{pmatrix} \quad (4)$$

Here, we use Lagrange multiplier method. Then we maximize Equation (5).

$$\max_{\mathbf{q}, \lambda} \left[\sum_p^{N_f} \left(\max_{axis_i} \left(\mathbf{R}(\mathbf{nf}_p) \cdot \mathbf{n}_{p,axis_i} \right) \right) \right] - \lambda (q_0^2 + q_1^2 + q_2^2 + q_3^2 - 1) \quad (5)$$

We solve this maximization by iterative algorithm, because $\mathbf{n}_{p,axis_i}$ depends on \mathbf{R} . Using the rotation matrix \mathbf{R} , we can obtain rotated target surface whose normal vectors' directions are same as x, y, z axes. Therefore, even if target surface with any attitude is input, our fitting algorithm can generate initial voxels whose normal vectors are similar to those of rotated target surface.

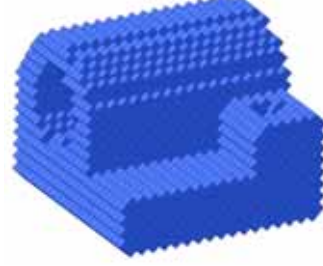


Figure 5: Initial voxels of joint.

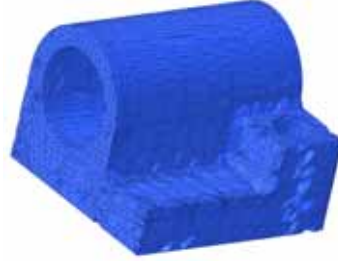


Figure 6: Failed mesh of joint.

4.2 Adaptive sampling

The Laplacian energy has a term of the squared norm of $\mathbf{v}_i - \mathbf{w}_i$, where \mathbf{v}_i is a boundary vertex of the surface and \mathbf{w}_i is its corresponding point on the target surface. Selection of the corresponding point is very important in this algorithm.

To represent sharp features, boundary vertices whose boundary edges corresponding to sharp features must be on the corresponding sharp features. So, it is important that \mathbf{v}_i whose \mathbf{w}_i is on a sharp feature is on the sharp feature.

In our previous research, we used up-sampling to increase the number of points on the target surface mesh in order to increase the number of candidates for the corresponding point. We used subdividing triangles for up-sampling. Let Q_j be the number of subdivisions of the boundary triangle t_j . We as-

summed the same value for Q_j , $j = 0, 1, 2, \dots$.

From experiments, we set Q_j to a large value in order to represent sharp features. There are, however, numerous sampling points. Thus, in this paper, we use adaptive sampling to decrease the number of sampling points. We increase Q_j whose t_j lie near the sharp features in order to increase only the sampling points near sharp features.

In this method, we check boundary triangle t_j whether near the sharp feature. Let D be threshold of the distance of boundary triangle t_j from nearest sharp feature.

Yellow points in Figure 7 are the points on the target surface obtained as a result of the new up-sampling. Using these candidate points, we fit voxels to the target surface.

5. RESULTS

In this section, we show some experimental results using our method. Figure 8 shows the target surface (joint model) after using our rotation algorithm. Figure 9 shows the boundary quadrilateral surface mesh of the all-hexahedral mesh (joint model) calculated by our new method. Figure 10, 11, 12 and 13 show the sampling points on the target surface (trim-star), initial voxels, the target surface after the rotation, and the result using by our method.

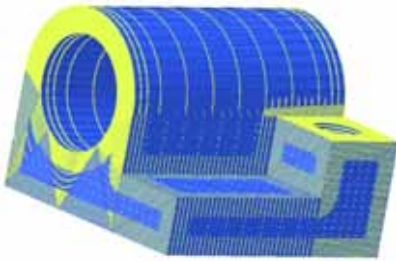


Figure 7: Up-sampling of the target surface mesh (joint).

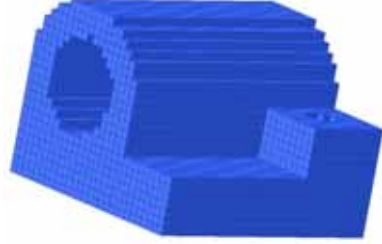


Figure 8: Voxels after using our rotation for target surface (joint).

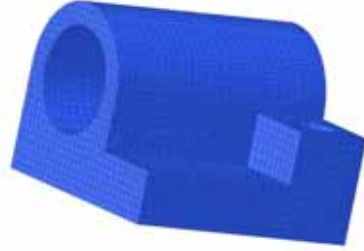


Figure 9: Result using by our method (joint).



Figure 10: Up-sampling of the target surface mesh (trim-star).



Figure 11: Initial voxels of trim-star.



Figure 12: Voxels after using our rotation for target surface (trim-star).



Figure 13: Result using by our method (trim-star).

6. CONCLUSION

In this paper, we explain our rotation and new up-sampling algorithms for target surface mesh, and show the experimental results of the proposed method. Using these algorithms for an automatic hexahedral mesh generation, we can obtain boundary surfaces of all hexahedral meshes having convex/concave sharp features.

ACKNOWLEDGMENTS

This work was supported by JSPS KAKENHI Grant Numbers 24760124 and 26420084.

REFERENCES

- [1] H. Date, S. Kanai, T. Kishinami, and I. Nishigaki. Mesh Simplification and Adaptive LOD for Finite Element mesh Generation. *Int.J. of CAD/CAM*. 6(1):No.8, 2006.
- [2] H. Date, S. Noguchi, M. Onosato, and S. Kanai. Flexible Control of Multimaterial Tetrahedra Mesh Properties by Using Multiresolution Techniques. *IEEE Trans. on Magnetics*. 45(3):1352-1355, 2009.
- [3] H. Kawaharada and H. Hiraoka. Boundary Stencils of Volume Subdivision for Simulations. In *Proceedings of the 2nd Asian Conference on Design and Digital Engineering (Shanghai, CHINA, Aug 27-29)*, pages 3-8. 2011.
- [4] H. Kawaharada and K. Sugihara. Hexahedral Mesh Generation Using Subdivision. *Computational Engineering*. 16(2):12-15, 2011.
- [5] Polymender [Internet]. [cited 2012 Jul 11]. Available from: <http://www.cs.wustl.edu/~taoju/code/polymender.htm>.
- [6] Y. Imai, S. Moriya, H. Hiraoka, and H. Kawaharada. Quadrilateral Mesh Fitting that Preserves Sharp features based on Multi-normals for Laplacian Energy. In *Proceedings of the 4th Asian Conference on Design and Digital Engineering (Seoul, KOREA, Aug 12-14)* 2013.

ABOUT THE AUTHORS

1. Yusuke Imai was born in 1992 in CHIBA, JAPAN. He is a master's student in the Department of Precision Mechanics, Chuo University. He is student member of Japan Society for Design Engineering (JSDE) and the Japan Society for Precision Engineering (JSPE). He received Bachelor degree in precision mechanics from the Chuo University in 2014.

2. Hiroyuki Hiraoka is a professor in the Department of Precision Mechanics, Chuo University. He has been active in the fields of life cycle engineering using information and communication technology. His research interests include reuse and remanufacturing using network agents and RFIDs, assembly using haptic interface, and product modeling. He is involved in the standardization activities of ISO TC184/SC4 on the representation of product model. Hiroyuki Hiraoka received Dr. Eng. Sci. degree in mechanical engineering from the University of Tokyo in 1983.

3. Hiroshi Kawaharada is an assistant professor in the Division of Systems Research, Faculty of Engineering, Yokohama National University. He has been active in the fields of Computer Aided Design and Computational Geometry. His research interests include automatic generations of surface and volume meshes. Hiroshi Kawaharada received Dr. Information Science and Technology degree from the University of Tokyo in 2007.

HISTORICAL PERSPECTIVES OF PERSPECTIVE

Adrian H. TAN, Fabian H. TAN, Frank M. CROFT Jr.

The Ohio State University, United States

ABSTRACT: For over 40,000 years, artists have been experimenting with the projection of three-dimensional objects onto surfaces by overlapping (layering) the objects, furnishing them with shadows, shrouding them with mist and clouds, and placing them on top of each other; yet the perceived depth we were searching for would come relatively recently, ca. 2,300 years ago, credited to the ancient Greeks. It had to be a ‘giant leap’ for them to foreshorten distant objects, such as to paint one’s arm smaller than the other, just because it was projected farther away from the same body. Then, after a hiatus, in the 15th century AD, foreshortening and linear perspective made a come-back after Brunelleschi repackaged them through mathematical computations. We are still using this concept in another attempt to merge the 2D and 3D environments in what we now call virtual reality.

Keywords: Foreshortening, history of art, perspective, three-dimension, virtual reality.

1. INTRODUCTION: THE QUEST FOR 3D IMAGES

One of the most ancient forms of art, which includes the origins of artistic principles still used today, is the projection of objects onto a two-dimensional medium. When early humans found that natural colorants, such as chalk, limestone, burned bones, red ochre, could be used to paint on media such as stones or cave walls and ceilings, they began to experiment with these new natural canvases. Among the earliest being the famous cave paintings in Altamira (in Spain) and Lascaux (in France). The Altamira cave paintings were datable to 40 – 35 ka while the Lascaux (Fig. 1) to 18 ka. They were likely painted by the Cro-Magnon or H. sapiens during the Upper Paleolithic Age.

As humanity progressed, these rock and cave paintings spread and could be seen almost everywhere in the world. One example is the Tham Khon prehistoric cave rock painting in Phu Phrabat (Fig. 2), in northern Thailand, ca. 3 to 2 ka, depicting several men walking in a line.



Figure 1: An auroch (*Bos primigenius*), Lascaux Cave paintings, ca. 18,000 years ago.



Figure 2: Rock painting in Phu Phrabat, Udon Thani, Thailand, ca. 2,000 to 3,000 years ago.

As the practice of paintings became more common to modern humans, almost any other

media could become a canvas. Examples are stelae (funerary stones), walls, ceilings, and clay-baked vessels. Fig. 3, for instance, shows a worship ritual of an Egyptian god by a family of five painted on a limestone stela (ca. 7th to 4th century BC during the Egyptian Dynasty Late Period).



Figure 3: A stela (funerary stone) made of limestone, Late Period, ca. 7th to 4th century BC (The Oriental Institute of the Univ. of Chicago).

An image of a male walking away from a piece of clayware was painted on the wall enclosure of a tomb (ca. 470 BC) in Paestum, in southern Italy (Fig. 4a) and a funerary game on a tomb wall, ca. 510 BC, in Tarquin, Italy (Fig. 4b). Fig. 5 illustrates more recent images of judgment days painted on the ceiling of the Kerta Gosa timber pavilion in the Island of Bali. Images were painted in Javanese *wayang* (shadow-puppet) style, while the story was cultivated based on the ancient Hindu lore in India. The Greeks had been especially prolific in painting their clayware with innumerable images as if they were story-books for children (and they often autographed them, just like today's book authors). One example is the depiction of a *hetaira* (courtesan) by the Painter Epiktatos on a wide-bowled drinking cup, ca. 510 BC, as shown in Fig. 6.



Figure 4: (a) left: A wall of the *Tomb of the Diver* in Paestum, ca. 470 BC (Paestum Archeological Museum, Paestum, Italy), (b) right: An underground wall of *Tomba dei Giocolieri* ca. 510 BC, Tarquin, Italy.



Figure 5: *Wayang* painting on the ceiling of Kerta Gosa timber pavilion, in Bali Island.



Figure 6: A Red-Figured *kylix* (wide-bowled drinking cup) by the Painter Epiktetos, ca. 510 BC (Museum of the Ancient Agora, Stoa of Attalos, Athens, Greece)

In light of this proliferation, one important question must be asked: *What do all these paintings* (Figs. 1 through 6) *have in common?* Careful observation will reveal that in each of

these paintings, the main object(s) were ‘twisted.’ In Figure 1 the horns and hooves of the *auroch* were somewhat twisted towards the painter. The same is true for the humans painted in Figures 2 through 6; they all experienced, in various degrees, what we will refer to as *perceptive distortion*. While their legs are marching or pointing towards left or right, their torsos are facing towards the viewer, as if the painters wished to show all parts of the torso and legs the same way they saw in real life, i.e., in front-view. The problem with this is that, in real life, the objects they saw were in three dimensional (3D) surroundings, but when they were painted, the media could only accommodate two-dimensional (2D) images. Hence, what they did was to display the upper part of the image (torso) in *front-view* and the lower part (waist and legs) in *side-view*, and thus created the notion of perceptive distortion. With respect to the *auroch* or other animals, they would see them in side-view, but seldom in front-view (unless they risked being stampeded); thus, they distorted the horns and hooves accordingly, such that the auroch is also painted in a combined *side-view* and *front-view*.

The quest for ‘three-dimensional’ images is not new. In what appears to be the first attempt to use a two dimensional medium to project a three-dimensional image, a shaman was painted on the roof of a cave in the Grotta di Fumane, near Verona, in Italy, probably by a Cro-Magnon or a Homo sapiens about 30 ka. A chunk of the roof containing the shaman image fell to the ground and is now preserved in the Sant’Anna d’Alfaedo Museo Paleontologico e Preistorico (Fig. 7). From the front view, the shaman looks like a simple red ochre “stick-man” painting on a flat surface; however, if it is turned to show its side-view, the piece of rock is not flat but rather has a ridge, onto which the shaman’s body was painted. It is likely that the painter attempt to draw the ‘torso’ of the stick-man protruding from a flat surface. If so, this could very well be the first ‘three-dimensional’ image known to humankind.



Figure 7: Front- view (left) and side-view (right) painting of a shaman, Fumane Cave, ca. 34 to 32 ka. (Sant’Anna d’Alfaedo Museo Paleontologico e Preistorico, 2010)

2. CARVING THE MEDIA

The stage was now set for a different approach to join in the human quest for realism in the two-dimensional media. Ancient artists began to carve the surface of the media, among the first being rocks. Aside from maintaining their permanence, probably the ancient artists had already the sense that to achieve a 3D surrounding they had to furnish the image with a *depth* that could be attained by sinking the image into the media. And thus, *petroglyphs* emerged in numerous parts of the world, such as in Valcamonica, in northern Italy, made by the ancient Camunnian tribes (Fig. 8), ca. Iron Age (first half of the first millennium BC) and in Wadi Rum, Jordan, by the Nabatean Arabs (Fig. 9), ca. 5 – 2 ka. Notice that, in Figure 8, the wheels of a wagon drawn by two horses engraved on the rock are represented by four *circles* as opposed to *ellipsoids*, commonly used to represent circles in a *pictorial* view (the simplest 3D view). The engravers placed these wheels and horses in a side-view, the position they would normally saw in real life, while the cart (represented by a rectangle) in a top-view, also the position they would normally see, i.e., from the top when they were on/in the wagon. Thus, Figure 11 displays a combined *side-view* and *top-view* of the wagon. In the Wadi Rum, the human is shown perceptively distorted in

both *side-view* and *front-view*, similar to human images painted in Figs. 2 through 6.

The provision of only two orthogonal views (front and side views or side and top views) would not allow our parietal lobe to perceive the visual and spatial relationships in a 3D environment. This is analogous to the use of only two instead of three orthogonal views (front, side, and top views) required for a graphic engineer to create a pictorial view of an object; or the use of only two of the three building views (plan, front-elevation, and side-elevation) for a contractor to build a house.

The most well-known perceptive distortion is those painted or engraved by the Egyptians, so much so, some would perceive it as a 'style' rather than an imperfection. Hence, modern day dancers (break dance) and pantomimists often adopted this Egyptian 'style' into their performances. In an interesting twist, when modern-day artists converted the images displayed in the Egyptian's *Book of the Dead* into figurines, they retained the distorted feature of the images and adopted them into the figurines as if to preserve the perceptive distortion as a 'style' (Fig. 10).



Figure 8: A petroglyph (rock engraving) of a four-wheeled cart in the Naquane Park, Valcamonica, Italy, ca. Iron Age (first half of the first millennium BC).



Figure 9: A petroglyph in Wadi Rum, Jordan, by the Nabatean Arabs, ca. 2,000 – 2,500 years ago.



Figure 10: Top: A part of the original Egyptian *Book of the Dead* (ca. 11th to 7th century BC); bottom: Distorted figurines made by modern-day artists (The Oriental Institute of the University of Chicago).

Achieving a *depth* to display the 3D effect could also be reached by an extrusion. Thus, the next attempts were made by carving the media with an extrusion, or *bas-relief*, such as those displayed by the grotesque figure raised from the surface of a Greek clayware from Athens in Figure 11, the courting couple in the wooden door in the Vatpa Phonphao in Luang-Prabang, Laos, in Figure 12, and the bas-relief marble slab of the chariot and horses in a race during the Panathenaic Games, in Athens, ca. 4th century BC, in Figure 13. In Figures 11 and 12, the images were once again carved as seen from both the *front-view* for the torso and *side-view* for the waist down, as if to repeat those painted on rocks and caves. In another effort to display the 3D effect, in Figure

13 the four horses were displayed with all four (countable) feet and overlapping torsos. The carvers believed (or wished to give the impression) that in so doing, the viewers could see all the animals as if in a real-life action. However, the horses, the wheel, the charioteer and the warrior were all depicted in an orthogonal side-view. Unless these horses were created for a viewpoint at a certain angle or unless the distance of each horse to the charioteer varies, the scene is technically unrealistic.

It was common for painters and carvers at the time to display animals with all of their legs visible along with overlapping torsos, as seen in real life. As an example, a tomb painting shown in Figure 14 also shows a chariot-racing theme, painted on a wall of the Andriuolo tomb, in Paestum, southern Italy, ca. 340-330 BC. In a similar notion, the horses were painted in an overlapping mode as if to give the impression that they were viewed at a certain angle. Still, except for the wheels, the scene was shown from orthogonal side-view. As compared with the *circular* wheel depicted in Figure 13, however, both wheels are shown as *ellipsoids*, as if they were made in a pictorial view. Despite the inconsistent direction of the horses, charioteer, and wheels, this painting combines *side-view* and *pictorial* settings. A much improved image of a chariot and the handlers (as if viewed from a certain angle) can be seen in Figure 15 in a mosaic assembly, ca. 2nd to early 1st century BC. Note also that the ellipsoid representing the spoked-wheel was done consistently.



Figure 11: Red clayware *lagynos* (wine or water jug) with grotesque figures applied to it, ca. 1st century BC to 1st century AD (Museum of the Ancient Agora, Athens, Greece).



Figure 12: A door in the Vatpa Phonphao in Luang-Prabang, Laos, shows carving of a hunting scene (right) and courting scene (left). Undated, the originals were probably made about 1,000 years ago.



Figure 13: A race in the Panathenaic Games, ca. 4th century BC (Museum of the Ancient Agora, Stoa of Attalos, Athens, Greece).



Figure 14: Four horses in a chariot-racing theme, Andriuolo tomb, ca. 340 -330 BC (Museo Archeologico Paestum, 2011)



Figure 15: A chariot and its occupants, ca. 2nd to early 1st century BC (Museo Archeologico Nazionale di Napoli, Italy).

The quest for 3D settings was still elusive at the time. Attempts at the third dimension on a two-dimensional medium, by overlapping objects were still at their infancy. *Sunken* or *bas-relief* would not suffice unless efforts were made for a *deep* or *high relief* images, where objects were sunk or raised at least a *half* of their actual depth from the base plane. Examples are the high-relief marble angel in a Siena church made during the Renaissance period (Fig. 16) and a deep-relief Buddha in Chongqing carved from rock during 12th to 13th century AD (Fig. 17). Notwithstanding, this approach was achieved by changing a 2D into a 3D medium, akin to what ancient humans did when they first created their Venus figurines, except that now the products could be made much bigger than before and undetached from the original material. So, the race to quench the thirst for creating 3D images on a 2D media continued.



Figure 16: A high-relief marble sculpture on the wall of a church in Siena, Italy made during the Renaissance era.



Figure 17: A deep-relief rock carving in Baodingshan Dazu Grotto, Chongqing Municipality, China, ca. 12th to 13th century AD.

3. OTHER ATTEMPTS

Some ancient artists began to experiment by blurring distant objects, employing shadows, mist or clouds to obscure distant objects, or placing these objects on top of the main figures. In an ancient painting of Europa riding a bull (Fig. 18), ca. 2nd to 1st century BC, the trees and mountain walls at the background were blurred so as to deemphasize their importance relative to the main objects and to promote the ‘extrusion’ of these objects from the background.

Numerous Chinese paintings or engravings had their objects shrouded by clouds. One example, shown in Figure 19 on a wall painting in the Mogao Caves in Dunhuang (an important trading center on the ancient Silk Road in China), displays ancient gods that were elevated on top of the clouds. Figure 20 shows two 2,300

years old Greek vessels with Red-Figured images painted by the Painter Spinazzo. A flying female figure was positioned on top to indicate its distance from the main figures.



Figure 18: Europa was seen here carried by a bull (Zeus in disguise) to Crete is entertained by her companions, ca. 2nd - 1st century BC (Museo Archeologico Nazionale di Napoli, Italy)



Figure 19: A wall painting in Mogao Caves in Dunhuang, China, ca. 10th century AD, displays the gods covered with clouds



Figure 20: Two Red-Figured Greek vessels by the Painter Spinazzo, Paestum, ca. 3rd century BC (Museo Archeologico Paestum, Paestum, Italy)

One more example is shown in Figure 21 depicting adobe brick making in Egypt's antiquity that can be seen in the tomb of Rekhmire, a noble from Thebes, ca. 3,500 years ago. The figure, reconstructed for clarity (original was disfigured), shows a water pond surrounded by trees on the left and brick workers hoeing the clay on bottom right and molding the brick on top right. Notice how the painter painted the trees and the worker carrying a water jug from the side-view but the water pond from the top-view. Notice also how the bricks were laid by the worker on top right as if they were stacked. They were not stacked, they were seen from the top-view but the worker was seen from the side-view. The top most bricks were assumed the farthest from the viewers; hence, they were placed on top. The painters' inability to create a *perceived depth* on the wall, had forced them to combine top and side views in an inconsistent way, similar to those presented in earlier figures.



Figure 21: A scene in adobe brick making in ancient Egypt (redrawn for clarity) painted on a tomb wall of Rekhmire in Thebes, Egypt, ca. 1500 BC.

4. IN SEARCH FOR DEPTH

The earliest type of pictorial view is probably the *oblique view*, which consists of a frontal view with depth being represented by parallel but angled features extruded behind it (Fig. 22a). With the oblique box, the front and back sides are the same size, and the edges representing the depth are slanted. The rear-view is analogous to the *shadow* of the front-view. This type of view incorporates “forced depth” into a two-dimensional viewpoint, making it easier to depict but still less convincing than the more advanced methods described next. Other types of pictorial settings are the *axonometric views*, which can be categorized into three different types: *isometric*, *diametric*, and *trimetric*. The corner and surrounding edges of the sample box can be used to define these three terms based on the perceived angles between each edge. *Isometric views* (Fig. 22b) have equal values (120°) for all three intersection angles at point A (red dot); *diametric views* (Fig. 22c) have two equal-valued angles and a different third angle; and *trimetric views* (Fig. 22d) have differing values for all three angles. The isometric view is a default in engineering graphics software since it yields the same scales in the x, y and z axes, and makes it easier to measure an object for its design and manufacturing processes. Nevertheless, in ancient times, no distinction appeared to have been made between oblique and axonometric views.

These pictorial views (chiefly oblique)

were introduced and brought by the Greeks to Italy as early as the 2nd to 1st century BC. Several mosaic images, such as *The Academy of Plato* (Fig. 23) from Pompeii, display background images of columns and an architrave showing oblique images. The two boxes at the bottom were also created in an oblique style. Another pictorial setting is shown in a marble relief of *Dionysus Visit to Icarus*, ca. 1st century AD (Fig. 24). The buildings behind the wall, especially the window on the right, reveal parallel lines that form an oblique view. While the oblique or axonometric pictorial views allow the perceived depth, they are still unrealistic representations of 3D space because the human eye perceives distant objects as being smaller and as such a different kind of viewing mechanism was needed.

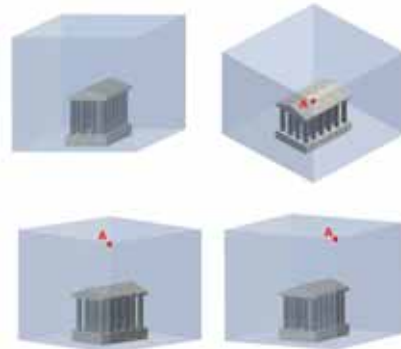


Figure 22 (top left clockwise): (a) oblique, (b) isometric, (c) diametric, and (d) trimetric pictorial views.



Figure 23: Mosaic of *The Academy of Plato*, Pompeii, ca. 2nd - 1st century BC (Museo Archeologico Nazionale di Napoli, Italy)



Figure 24: A marble relief of *Dionysus Visit to Icarus* ca. the 1st century AD (Museo Archeologico Nazionale di Napoli, Italy).

5. PERSPECTIVES

A 15th century artist, engineer and architect, Filippo Brunelleschi (responsible for building *il Duomo*, the world's largest brick and mortar dome) in Florence, Italy, is credited with the development of geometry-based *linear perspective*, another type of pictorial view, in which viewing lines converge on a vanishing point on the horizon line instead of being rendered in parallel. Brunelleschi computed the foreshortened dimension of a distant object through geometry. Several schemes of linear perspective exist, with a varying number of one, two, or three vanishing points as shown in Fig-

ure 25.

A perspective requires a feature called *foreshortening* that would render distant objects smaller in size and dimensions. This trait was discovered much earlier than Brunelleschi's linear perspective; the ancient Greeks discovered it over 2,300 years ago. One of the most primordial perspectives is shown in Figure 26 of the Apulian wedding scene by the Painter Baltimore, in which the foreshortening, subtle as it is, is evidenced from the smaller right (distant) breasts of the man and woman and the smaller right arm of the flying woman on top. A more distinct specimen is the small shrine known as an *aedicules*, made from cuts of marble stones called *opus sectile*, depicting a drunken Silenus (Fig. 27). Notice how the foreshortening took place on the roofs on top of the male and female deities next to Silenus, giving the impression of depth.

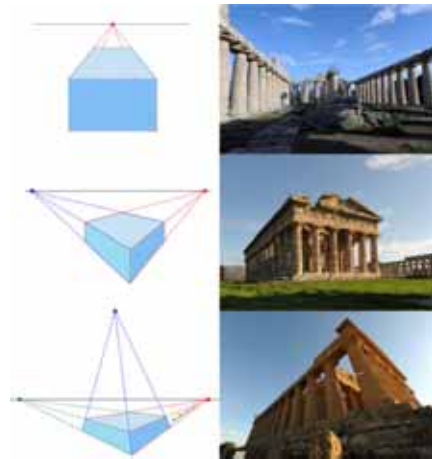


Figure 25 (top-down): Perspectives of Greek temples in Paestum, Italy, in photos as we see them in real life, with one, two, and three vanishing points at the horizon.



Figure 26: This Apulian Red-Figured genre shows a wedding scene, by the Painter Baltimora, on a ceramic dish, ca. 330 BC (Villa Giulia National Etruscan Museum, Rome, 2010).



Figure 27: An *aedicules* (small shrine) made from *opus sectile* depicting drunken Silenus (companion of Dionysus) assisted by satyrs, ca. 1st century AD (Museo Archeologico Nazionale di Napoli, Italy).

Prompted by Brunelleschi, linear perspective flourished in the 15th century AD. Figure 28 shows Fra Angelico's *Annunciation* painting ca. AD 1436. The famous painter made foreshortened the man, woman and the angel on the left background as well as the Ionic capitals of columns at the back rows through linear perspective. This practice was to be further developed in the 16th century – the German painter Albrecht Dürer worked with single-point perspective considerably when illustrating his 1511 text, *The Life of the Virgin*, using the aid of window-like “perspective machines” (Fig. 29) to master and experiment with the tech-

nique [6]. From then on, perspective became a daily staple of painters all across the world.



Figure 28: *Annunciation* by Fra Angelico, ca. AD 1436 (Prado Museum, Madrid, Spain).



Figure 29: *An Artist Drawing a Seated Man* by Albrecht Dürer, ca. AD 1525 (The British Museum, London, United Kingdom).

6. 3D PERSPECTIVES

What was called 3-D in not-too-distant past has now become ‘2½-D’ or ‘pseudo 3-D’, which unlike 2-D and 3-D could not be represented in mathematical forms. The 2½-D is a representation of a 3-D object in a 2-D platform. Hence, just adding a certain depth to a 2-D image nowadays does not qualify an object to be considered as a 3-D object. The 2½-D is often used to represent the axonometric drawings [5]. Another term employed for 2½-D with a perspec-

tive taste is '¾-perspective' or '¾ view' which is similar to axonometric or perspective with two-vanishing points. In today's world, a 3-D model encompasses not just visual but also mathematical and computing techniques. Three-dimensional perspective in graphics is used in a number of different applications, such as object modeling and display, environment reconstruction and simulation, and scene projection.

In the latest quest for reality, modern humans seem to settle with an interactive virtual reality (VR) environment and its derivatives. Several VR models now proliferated in numerous domains, including in engineering and education [2]. By blocking the eyes and ears with instruments that display 3-D objects and environment, visual and audio related brain lobe activities are cut off from the real surroundings. Images are transferred into a virtual space to allow human perception immersed in a different world. Imagine a present-day human going back in time 18 ka, entering a cave and virtually interacting with the Cro-Magnon; they may probably paint the cave walls together with them. However, a drawback still exists in that to immerse in the world of virtual reality one needs to be equipped with virtual reality instruments [1, 3-4].

7. CONCLUSION

The discovery of the concept of foreshortening by the ancient Greeks and later spread to southern Italy, about 2,300 years ago would probably be the culmination of human achievements in projecting real objects onto flat surfaces. Since time immemorial, we would naturally approach a distant object in order to verify and clarify what it was (now we can use optical instruments). However, what took us so long was likely related to the difficulty of our parietal lobe to conduct spatial association in scaling down images when we projected them onto a flat surface. It would be unimaginable for us in ancient time to paint or carve a distant person smaller than the other, especially since they

were about the same size when we met them in person. Now of course our conviction of foreshortening had been greatly help by photographs and thus, it would now be difficult for us to accept, as realistic, say in a picture of two people having the same dimension while one of them is photographed at a distance.

Another interesting historical observation is that after the invention of the concepts of foreshortening around 330 BC (e.g., in vase paintings), of oblique images (e.g., in mosaics) around 2nd to 1st century BC, and of linear perspectives around the 1st century AD (e.g., in *opus sectile*), the spread and use of approaches seemed to halt and began to flourish again during the Renaissance in the 15th century AD, primarily in Europe.

REFERENCES

- [1] Hadipriono, Fabian (Tan). "Virtual Reality Application in Civil Engineering," pages 93-100, 196-197. ACM Symposium on Virtual Reality Software and Technology, Hong Kong, 1996.
- [2] Hadipriono, Fabian (Tan) and Ashraf S. Barsoum. "Modeling for safety against falls from form scaffolding in a virtual environment," pages 119-139. Civil engineering and environmental systems, Colchester, UK: Taylor & Francis, Vol. 19, No. 2, 2002.
- [3] Hadipriono, Fabian (Tan), Barsoum, A.S., Tsay, T.C., Nemeth, Z.A., and Larew, R.E. "Visualization and Graphics in INTREPID-VR," pages 107-113. Proceedings of the second International Conference on Visualization and Intelligent Design in Engineering and Architecture (VIDEA 95), La Coruna, Spain, June 1995.
- [4] Hadipriono, Fabian (Tan), Richard E. Larew, Tsung-chieh Tsay. "The Ohio State University Interactive and Immersive Training in a Virtual Environment for Construction Students." Session 2526, 1996 ASEE Annual Conference Proceedings, 1996.

- [5] Marr, D. and T. Poggio. 1979. "A Computational Theory of Human Stereo Vision," pages 301-328. Proceedings of the Royal Society of London. Series B, Biological Sciences, Vol. 204, No. 1156 (May 23, 1979).
- [6] Price, David. *Albrecht Dürer's Renaissance: Humanism, Reformation, and the Art of Faith*, pages 146-147. University of Michigan Press, 2003.

Environmental Engineering and Geodetic Science at The Ohio State University, where he has been since 1984. Dr. Croft has a B.S. in Aerospace Engineering from the Indiana Institute of Technology, an M.S. in Engineering from the West Virginia College of Graduate Studies, and a Ph.D. in engineering at Clemson University in South Carolina. Dr. Croft has been very active with the Engineering Design Graphics Division (EDGD) of the American Society for Engineering Education (ASEE) since 1974.

ABOUT THE AUTHORS

1. Adrian H. Tan is a graduate student at the Ohio State University. Adrian has a B.S. in Computer Science and Engineering and an M.S. in Civil Engineering from the Ohio State University. Adrian is currently working towards a Ph.D. in civil engineering and construction with a focus on computer graphics and virtual simulation in the engineering industry.
2. Dr. Fabian H. Tan is a Professor in the Department of Civil and Environmental Engineering and Geodetic Science at The Ohio State University, where he has been since 1982. Dr. Tan has an M.S. in structural engineering, an M.E. in construction engineering and management, and a Ph.D. in construction engineering and management from the University of California in Berkley. Dr. Tan has served as a forensic consultant with numerous law firms in Ohio as well as the United States Air Force Weapons Laboratory in Kirtland, New Mexico. Additionally, Dr. Tan has also served as a project manager and engineer on multiple international construction projects such as Leighton Australia-Indonesia Construction Inc. and Mahkota-Ekman Sweden Inc. Dr. Tan is currently teaching History of Ancient Engineering at The Ohio State University.
3. Dr. Frank M. Croft Jr., P.E., is an Associate Professor in the Department of Civil and

HOW TO PROMOTE STUDENT CREATIVITY AND LEARNING USING TUTORIALS IN TEACHING GRAPHICS AND VISUALISATION

Anders HAST

Uppsala University, Sweden

ABSTRACT: Course assignments play an important role in the learning process. However, they can be constructed in such a way that they prohibit creativity, rather than promoting it. Therefore it was investigated how programming assignments are set up that students encounter in computer science education and which approach could help the students in problem solving and whether these would help or prohibit them to be creative or not. Especially, an online tutorial about visualisation using VTK and Python was used as an example in different courses on visualisation. It was also examined how students in the computer graphics courses that did not have access to such tutorial answered questions about assignments.

Keywords: Tutorials, Creativity, Motivation, Assignments, Scientific Visualisation and Computer Graphics programming

1. INTRODUCTION

Course assignments play an important role in the learning process and often allow the students to comprehend the theoretical part of the course by applying it in practice. However, programming assignments in general, also those in graphics and visualisation, can be constructed in such a way that they prohibit creativity. Instead, the task is merely limited to find out what the teacher expects, rather than actually promoting the learning process, the students enthusiasm and the self confidence that comes from overcoming problems that are actually understood. Therefore it was investigated how programming assignments are set up that students encounter and which approach would be most beneficial for them. The idea was to examine, in particular, to what degree online tutorials could help the students in problem solving and whether these would help or prohibit them to be creative or not. An online tutorial about visualisation using VTK and Python was used as an example in different courses on visualisation. The students had access to it during the assignments and it explains how to make

scientific visualisations and what basic graphics algorithms that are needed. The task was to solve problems that were described with pointers to the tutorial so that they could find the essential program examples, which would help them to solve the task. It was also examined how students in the computer graphics courses that did not have access to such tutorial answered questions about assignments.

2. EVALUATING ASSIGNMENTS IN COMPUTER SCIENCE EDUCATION

Assignments in computer science education (CSE) can be constructed in many ways. However, three different approaches for constructing assignments were identified at forehand:

- A. A code skeleton is provided where the teacher has removed some lines from a working program. The students task is to fill in the missing lines according to some specification.
- B. A working program is provided. The students task is to add new functionality to that program according to a specification.

- C. A tutorial with more or less complete programs describing different functionality is provided. Your task is to write a program according to a specification. The students are free to use any program or functionality from the tutorial.

The courses involved in the evaluation presented in this paper used to have predominantly assignments of type **A**. The idea was that creativity would be notably promoted by changing them to type **B** or **C**. The hypothesis was that too many courses still have type **A** and that students would prefer to have assignments of type **B** and **C**. Especially, it was interesting to find out how type **C** assignments were perceived by the students. Subsequently, the computer graphics course was changed to have assignments of type **B** and the scientific visualisation course was changed to have type **C**. All assignments are found online at studentportalen [4]. Most of them contain a lot of explaining text, images and suggestions, so they are in fact a mini tutorial by them selves.

The computer graphics course [1], which gives 10 credits, uses the textbook by Angel and Shreiner [8] and 23 students in total answered the questionnaire. The students have to do three assignments and one larger project.

The scientific visualisation course, [2] gives 5 credits and uses the textbook by Schroeder et al. [20]. Even if the assignments have been changed to be of type **C**, some have pointers to working code examples. Nevertheless it is expected that the students use the tutorial [6] to figure out how to do the required changes defined in the assignments. A total of 24 undergraduate students answered the questionnaire. An example page from the tutorial is shown in Figure 1, where it is explained how conversion of files to VTK format can be done.

The graduate course on scientific visualisation has been given in two variants and a total of 16 students answered the questionnaire. The first was given as a two day workshop by UPPMAX [7] and the second was a graduate course given by SeSE [3], which have a syllabus very much

like the under graduate course. However, the project is aimed at visualising the data that the PhD students use in their own research. UPPMAX is the Uppsala Multidisciplinary Center for Advanced Computational Science, which is a High performance computing (HPC) resource at Uppsala University. UPPMAX is one of six High Performance Computer centres in Sweden under SNIC [5], which is the Swedish National Infrastructure for Computing. SeSE is a Swedish research school aiming at fostering collaboration between Swedish PhD students in different fields and giving courses in the field of HPC, visualisation, big data and other related areas.

A questionnaire was filled in by the students and the results from the students answers will be discussed. Moreover, conclusions will be drawn about how to proceed in the ongoing work to improve assignments in order to promote student creativity and learning.

3. CREATIVITY, MOTIVATION AND BUILDING NEW KNOWLEDGE

Assignments can be constructed in such a way that the students feel that they have to figure out what the teacher intended and follow his line of thoughts, rather than being creative, exploring and getting new insight. This kind of assignments is typically not motivating for the students to learn more than just doing the task given them. Type **A** is a typical example of such assignments. The other types allow students and motivate them to use their gained skills and new knowledge to explore and discover further. By making use of the tutorial they are, in an indirect way, encouraged to also add more functionality than requested in the assignment. This is often easy to do in graphics and visualisation as the result is visual. By changing colours as well as appearance of objects, glyphs, or whatever is on the screen, the students can play around with the application and change it to their liking.

Creativity can be promoted in many ways. In both courses it was provided extra credit opportunities, as discussed by Roberts [18]. Furthermore,

Converting Data Files

Linked from [VTK Tutorial](#), [Reading Files](#)

If you would like to convert the data you have into the `vtk` file format so that you can use your favorite `vtk` program to visualize it, then you can use a writer in `vtk` that will do the job for you. This short tutorial will show you how easy it is to convert data, in this case the data files used in [How to Read Text Files](#). Please download the code [Convert.py](#). Don't forget to give the file names as an argument when you run the code:

```
python Convert.py data1.txt data2.txt data3.txt new_data.vtk
```

Contents [hide]

1 Explanation of the Code

2 Exercises

2.1 Visualization

2.2 Convert from OBJ to VTK

Explanation of the Code

Load all modules.

```
import sys
from vtk import *
```

Import the function `readPoints` from the `ReadPoints.py` file explained in: [Reader for ASCII Files](#).

```
from ReadPoints import *
```

Create the unstructured grid and use your reader from the previous exercise.

```
data=vtk.vtkUnstructuredGrid()
data.SetPoints(readPoints(sys.argv[1]))
data.GetPointData().SetVectors(readVectors(sys.argv[2]))
data.GetPointData().SetScalars(readScalars(sys.argv[3]))
```

Use a Unstructured Grid Writer to write the data to the `.vtk` file. Yes, that is all you need to do!

```
Data=vtkUnstructuredGridWriter()
Data.SetInput(data)
Data.SetFileName(sys.argv[4])
Data.Write()
```

Exercises

Do the following exercises!

Visualization

Change the program [BallsArrows.py](#) so it will use your `vtk` file instead of reading the individual textfiles.

Convert from OBJ to VTK

Make a program that converts a `.obj` file to `.vtk`. Change the program [Teapot.py](#) so it will read the `vtk` file instead. You can actually use the [vtkDataWriter](#), which is the parent class of [vtkUnstructuredGridWriter](#).

Figure 1: Screenshot from one web page in the VTK Tutorial that describes how to convert data ASCII files produced by Matlab® or any other program to standard VTK format. The code is explained in detail as well as how to run it. There are links to the source code and other pages explaining further details. Moreover, there are usually exercises and sometimes also answers.

Romeike [19] identify three drivers for creativity in computer science education, namely: the person with his motivation and interest, the IT environment and the subject of software design itself. Obviously, they are all interconnected and the aim was to improve the third driver by providing challenging and interesting tasks to perform, i.e. design an application that did something interesting.

It was noted by Knobelsdorf and Romeike [15] that "The activities mostly start with gaming, followed by exploring applications and their possibilities, experimenting with the computer, knowledge gathering, and the Internet usage." This observation was based on work by Knobelsdorf and Schulte [16], where it is noted that students starts as users and then become designers, which is a natural development from simple to more complex activities. On-line tutorials typically allow students to do this as they can explore the tutorial and add functionality according to their interest and not only because the teacher say so.

The advantage with such e-learning materials is that students can build their own conceptual understanding and improve their skills. Faessler et al. [13] discuss the advantages of PBL based e-learning material. Nevertheless, Boyer et al. [10] point out that many have investigated the cognitive aspects [11] of e-learning, rather than motivational aspects. Obviously, motivation is important as the first driver identified by Romeike.

Furthermore, tutorials are important as they present models of how things can be done, i.e. solutions to simple problems or general cases that can be applied by the students on their particular problem. VanDeGrift [21] discuss different learning theories and what helps students to learn. She points out that students build knowledge by looking at such prototype models [17]. For courses that include programming this would typically be pseudo code or working code examples provided by the teacher. Similarly, the constructivist learning theory [9, 12], claims that students build knowledge based on existing knowledge and upon the foundation of previous learning [14].

VanDeGrift say that "students must actively construct new ideas and experiment with those new ideas and mold them into existing knowledge".

4. RESULTS OF THE QUESTIONNAIRES

A questionnaire was handed out during one course on computer graphics for undergraduate students, one course in scientific visualisation, also for under graduates and also in yet another scientific visualisation course, but this time for graduate students. The questions asked were:

1. Which approach has been most common in programming courses you have taken?
2. Which approach do you prefer?
3. Which approach do you think will be most helpful for you to learn new things?
4. Which approach will allow you to be creative and try to invent new things?

The result of question 1 is illustrated in Figure 2. It is obvious that many students still do encounter questions of type A, just as was expected.

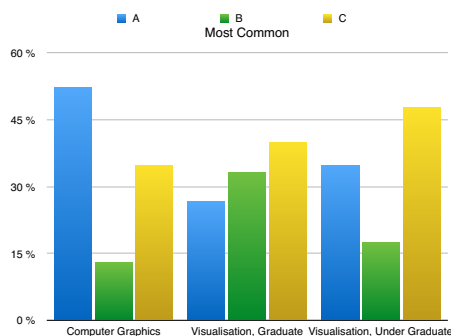


Figure 2: Outcome of the question: "Which approach has been most common in programming courses you have taken?"

The result of question 2 is shown in Figure 3 and it can be noted that many computer graphics students actually prefer type A, but for the

scientific visualisation course the types **B** and **C** are preferred by most students, while just a few advocate type **A**.

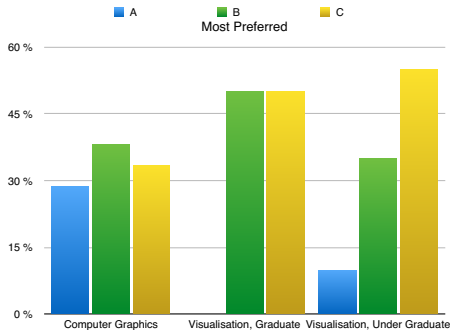


Figure 3: Outcome of the question: "Which approach do you prefer?"

When asking which approach they think will be most helpful for learning new things (question 3), there are some notable differences, as shown in Figure 4. Even if type **A** was preferred by the computer graphics students, they did not generally think it would help them learn new things. Likewise, type **C** was the most preferred in the scientific visualisation courses.

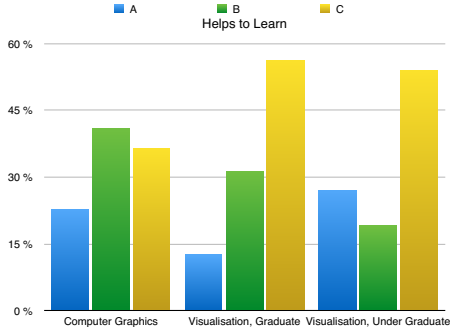


Figure 4: Outcome of the question: "Which approach do you think will be most helpful for you to learn new things?"

Even more computer graphics students realised that type **A** do not promote creativity. And once

again the scientific visualisation students preferred type **C**, as can be seen in Figure 5, where the result of question 4 is shown in the graph.

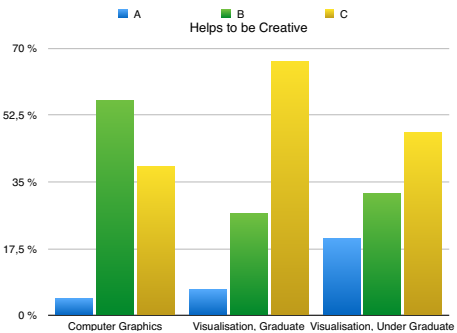


Figure 5: Outcome of the question: "Which approach will allow you to be creative and try to invent new things?"

Interestingly, more computer graphics students prefer type **B**, over **C**. For the scientific visualisation students, on the other hand, the opposite is true. Even though the number of students asked does not allow us to draw general conclusions, this difference might depend on the fact that the computer graphics students encountered type **B** during the course, while the scientific visualisation students encountered type **C**.

5. DISCUSSION

In retrospect it was understood that it was unclear for some students what the differences between the types really were. The idea was that type **A** was not a working program since code was removed from it. Nevertheless, some students seemed to have thought it was the same as type **B**. Hence, some choose **A**, when they were supposed to choose **B** and vice versa. There is also a risk that students in the visualisation courses confused type **B** with **C** as some assignments gave pointers to a working program, even if the tutorial always was pointed out as the main source of information.

Moreover, the number of students answering the questionnaire was rather low (23, 16, 24, for

each group respectively). Hence, one should not draw decisive conclusions by looking at the bar graphs. Nevertheless, the trend is obvious, also from the authors own experience, that type **A** assignments are still quite common. Furthermore, it is obvious that type **B** and **C** are generally preferred compared to type **A**.

6. CONCLUSION

The way assignments are designed can either prohibit or promote creativity. Still, many assignments contain non working code, where some lines are missing and the students are supposed to fill in that code. This type of assignments hardly allows students to be creative, instead they try to figure out what the teacher had removed. Moreover, they are not explicitly motivating the students to do more than just filling in the missing code.

Assignments based on working code, on the other hand, allows the students to explore the behaviour before adding new functionality. Playing around with the code, gives an initial understanding of what the code does and this is an important first step in the creative process. Then it is up to the student how to solve the task, given in the assignment. By having access to an online tutorial, the students can learn about how to solve such task in a generic way. They must themselves figure out how those solutions apply to their code. This might not promote creativity in itself. However, by reading through the examples and having access to exercises and solutions, the students will see what the possibilities are and hopefully they will be stimulated to try out things that they think can improve their code. This is especially true for assignments that produce visual results, which is always true for computer graphics and scientific visualisation.

It is interesting to note from the evaluation that the computer graphics students, which had not access to a tutorial in their course, generally were more positive to assignments that do not include a tutorial. While, on the other hand, the students in the scientific visualisation course, who had such

tutorial, were generally more positive to using tutorials than not having access to them. Therefore, it seems obvious to make the conclusion that a tutorial will not only help students to be creative and motivate them to investigate further, but a good tutorial will also make the students more positive to using it.

The aim is that the tutorial shall be extended and improved so that it can be an even more valuable tool for the visualisation courses and workshops given within SeSE and UPPMAX.

REFERENCES

- [1] Syllabus, uppsala university, 2010. URL <http://www.uu.se/en/admissions/master/selma/kursplan/?kKod=1TD388>. Accessed June 14, 2014.
- [2] Syllabus, uppsala university, 2010. URL <http://www.uu.se/en/admissions/master/selma/kursplan/?kKod=1TD389>. Accessed June 14, 2014.
- [3] Swedish escience education, 2013. URL <http://sese.nu>. Accessed June 14, 2014.
- [4] Studentportalen, uppsala university, 2014. URL <https://studentportalen.uu.se/portal/portal/uusp>. Accessed June 14, 2014.
- [5] Swedish national infrastructure for computing, 2014. URL <http://www.snica.se>. Accessed June 14, 2014.
- [6] Uppmax, vtk tutorial, 2014. URL http://www.uppmax.uu.se/docs/w/index.php/VTK_Tutorial. Accessed June 14, 2014.
- [7] Uppsala multidisciplinary center for advanced computational science, 2014. URL <http://www.uppmax.uu.se>. Accessed June 14, 2014.

- [8] E. Angel and D. Shreiner. *Interactive Computer Graphics: A Top-Down Approach with Shader-Based OpenGL*. Addison-Wesley Publishing Company, USA, 6th edition, 2011. ISBN 0132545233, 9780132545235.
- [9] M. Ben-Ari. Constructivism in computer science education. *SIGCSE Bull.*, 30(1): 257–261, March 1998. ISSN 0097-8418. URL <http://dx.doi.org/10.1145/274790.274308>.
- [10] K. E. E. Boyer, R. Phillips, and J. L. Michael Wallis and Mladen Vouk. Investigating the role of student motivation in computer science education through one-on-one tutoring. *Computer Science Education*, 19(2): 111–135, 2009. URL <http://dx.doi.org/10.1080/08993400902937584>.
- [11] R. De Villiers. Usability evaluation of an e-learning tutorial: Criteria, questions and case study. In *Proceedings of the 2004 Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists on IT Research in Developing Countries*, SAICSIT '04, pages 284–291. South African Institute for Computer Scientists and Information Technologists, Republic of South Africa, 2004. URL <http://dl.acm.org/citation.cfm?id=1035053.1035092>.
- [12] R. Duit. The constructivist view: A fashionable and fruitful paradigm for science education research and practice. In L. Steffe and J. Gale, editors, *Constructivism in Education*, chapter 14, pages 271–285. Taylor & Francis, 2012. ISBN 9781136476082. URL <http://books.google.se/books?id=f5Up2BpqdSsC>.
- [13] L. Faessler, H. Hinterberger, M. Dahinden, and M. Wyss. Evaluating student motivation in constructivist, problem based introductory computer science courses. In *E-Learn 2006, World Conference on E-Learning in Corporate, Government, Healthcare, & Higher Education*, pages 1178–1185. 2006.
- [14] W. A. Hoover. The practice implications of constructivism. *SEDL Letter*, IX(3), 1996.
- [15] M. Knobelsdorf and R. Romeike. Creativity as a pathway to computer science. *SIGCSE Bull.*, 40(3): 286–290, 2008. ISSN 0097-8418. URL <http://dx.doi.org/http://doi.acm.org/10.1145/1597849.1384347>.
- [16] M. Knobelsdorf and C. Schulte. Computer science in context - pathways to computer science. In R. Lister and Simon, editors, *Seventh Baltic Sea Conference on Computing Education Research (Koli Calling 2007)*, volume 88 of *CRPIT*, pages 65–76. ACS, Koli National Park, Finland, 2007.
- [17] D. A. Norman. What goes on in the mind of the learner. *New Directions for Teaching and Learning*, 1980(2): 37–49, 1980. ISSN 1536-0768. URL <http://dx.doi.org/10.1002/tl.37219800205>.
- [18] E. Roberts. Strategies for encouraging individual achievement in introductory computer science courses. *SIGCSE Bull.*, 32(1): 295–299, March 2000. ISSN 0097-8418. URL <http://dx.doi.org/10.1145/331795.331873>.
- [19] R. Romeike. Towards students' motivation and interest: Teaching tips for applying creativity. In *Proceedings of the 8th International Conference on Computing Education Research*, Koli '08, pages 113–114. ACM, New York, NY, USA, 2008. ISBN 978-1-60558-385-3. URL <http://dx.doi.org/10.1145/1595356.1595380>.

- [20] W. Schroeder, K. M. Martin, and W. E. Lorensen. *The Visualization Toolkit (4:th Ed.): An Object-oriented Approach to 3D Graphics*. Pearson Education, Inc., 2006. ISBN 1-930934-19-X.
- [21] T. Vande-grift. Encouraging creativity in introductory computer science programming assignments. In *ASEE, 2007 Annual Conference & Exposition*. 2007. Session: Programming for Engineering Students.

ABOUT THE AUTHOR

Anders Hast is associate professor at the division of visual information and interaction, department of information technology, Uppsala University. He is also an application expert in scientific visualisation at UPPMAX, Uppsala. His research interests are computer graphics, computer vision and visualisation. He teaches courses in computer graphics and visualisation and also regularly gives workshops in visualisation for both UPPMAX and the Swedish eScience Education (SeSE). Personal web page at the university: <http://www.cb.uu.se/~aht/>

HULL LINES PLAN. HISTORY. COMPARISON OF METHODOLOGIES FOR 3D MODEL GENERATION

Fernando FADON, Enrique CERON, Raul PEREDA, Laida FADON
Cantabria University, Spain

ABSTRACT: Definition of hull drawings requires specific methods as ships are made of complex geometries. The plan of lines is a particular form for representing hulls and is based on Descriptive Geometry developed by Gaspard Monge during the XVIII century. This paper provides a brief history of the plans of lines, as well an explanation of how measurements of ships were developed for later performances or verifications. Moreover, it suggests and explains how total stations, laser scanner or software can be applied for the generation of 3D models.

Keywords: lines plan, measurement, total station, laser scanner, hull model.

1. INTRODUCTION

Warped surfaces have been remarkably developed as their geometrical definition has improved due to mathematical formulas and virtual representations with specific software. Those shapes have developed for industry sectors such as automotive, aeronautic, naval or even in electrical appliances and other household tools.

Previously, modeling of these kinds of surfaces was necessarily accomplished by using scale models or graphical representations. The accuracy achieved in these methods was much lower than the obtained in other mechanical elements (cylinders, connecting rods, gears ...). Consequently, repeatability and verification of products became more difficult. Shipbuilding has been historically one of the first sectors which have had to cope with this problem, first with the development of ship designs and then with propellers. Another area where warped surface monitoring becomes relevant is topography.

This paper describes at first the way to represent the hull of a ship. Then, it includes a historical review explaining the evolution of those hull lines planes from XVI century to nowadays.

Furthermore, a measure process (equipment and method) has been developed in order to construct the hull based on the planes, and in the other way around, drawing the planes by capturing spatial data of the ship. 3D model generating procedure with a total station and a laser scanner is explained for the plane measuring and drawing processes. It explains the methodology applied, which makes easier and improves this process, as well as opening new ways for other different applications.

2. HULL LINES PLAN

The hull lines plan refers to the plane where the ship shapes get defined based on cross-sections which are parallel to the three orthogonal main planes (figure 1).

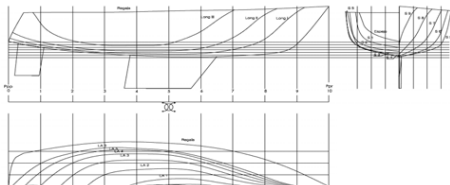


Figure 1: Hull lines plan.

On the front view, the left side represents the hull sections from the central rib to the stern, whereas the right side shows the ones from the central rib to the bow. The hull is divided in 10 or 20 cross-sections, depending on the kind of ship represented, as there is not a fixed rule for all of them. On the top view, water lines (horizontal hull sections) are displayed. Nowadays, the floating line coincides with the water line number 6 or 10 depending on the ship draught, where zero is the reference plane. It is represented only in the half of top view. On the other half of the top view oblique sections are drawn, which help to best define the shape of the hull. On the lateral view appear 3 to 6 longitudinal sections, which are distributed along half beam.

3. HISTORICAL REVIEW

The following lines briefly explain the evolution of Hull Lines representation and the representation systems. Orthographical projections have been applied since the Sumerian (3000 bC) and the Egyptian culture. But it is in the XVII and XVIII centuries when there is evidence of an evolution in posing and solving graphical and geometrical problems. Especially relevant is the multiview orthographic projection, explained in "Geometrie Descriptive" by Gaspard Monge in 1795. Through this system he solves problems such as those related with defiladed fortifications, which become relevant until the First World War. Perspective representation systems had a significant progress in the Renaissance.

However, the naval sector have often had to cope with geometrical construction problems and solve them in order to define the ship lines,

such as ribs and timbers, bow and stern shapes, the rising line over the keel and the hull. Through the orthographical projection, problems related to the hull representation have been solved. A selective bibliography of documents from that century with significant aspects showing the technic has been founded.

3.1 S XVI and XVII centuries.

Conserved documents from XVI and XVII centuries are limited and contain just a few figures.

F. Oliveira shows in his manuscript "O Livro da Fábrica das Naus" some drawings (Figure 2) in which are represented the most important characteristics for a ship construction. He applies the plane geometry and shows the folded rib for its representation.

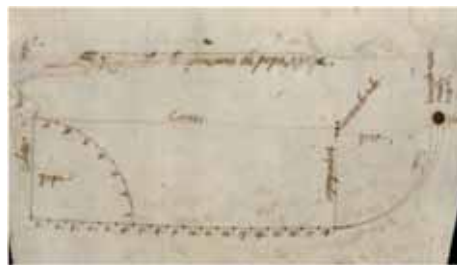


Figure 2: Oliveira (1580)

Looking at the drawings extracted from the manuscript of Matthew Baker "Fragments of Ancient English Shipwrightry" (1586) [2] (figure 3), some similarities with Oliveira can be appreciated, although they probably met each other as Matthew Baker lived in England during that period. In the illustrations he applies in an elementary way the orthogonal views, and in

the top view he draws the breadth line by taking from the ribs the point with biggest beam [3] In the other picture, the fish sketched in the underwater part of the hull, with its stumpy head and long, tapering tail symbolizes the ideal shape of hull gliding through the water.

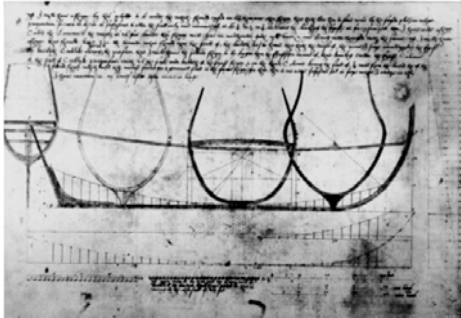


Figure 3: Design by the English shipbuilder Matthew Baker (1586)

“Deane’s Doctrine of Naval Architecture” (1670) [4] shows some works which are more elaborated. It is remarkable the use of orthogonal views, the application and procurement of water lines and the use of the third projection of rib box (figure 4). Moreover, it is also relevant the calculation process of the volume of the dived part of the ship, whereby the waterline and other buoyancy parameters.

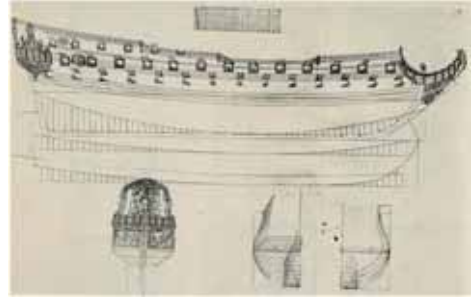


Figure 4: Chapter “The ship’s draught completed in every part” where Deane explained it.

3.2 Eighteenth century.

In the eighteenth century takes place important improvements. Specially significant is “Traité du navire” de Bouguer 1746 [5], who founded a discipline that defined not the rules for building a ship but the theories and tools to predict a ship's characteristics and performance before it was built. In Ships and Science, Ferreira [6] argues that the birth of naval architecture formed an integral part of the Scientific Revolution.

Mungo Murray explains that his book “A Treatise on Ship-building and Navigation” (1754) [7] is structured “In three parts wherein the theory, practice, and application of all the necessary instruments are perspicuously handled”. ... Remarkably interesting is the second part of the treatise (page 117) “Of the Orthographic projection of Solids on a Plane”. Here it is explained:

“The chief design of delineating a house, ship, or any other solid upon a plane, is to settle the just dimensions, and symetry of its parts according to the scheme of the builder. When this is done by mathematical rules, we can find the exact length, breadth and height, not only of the whole, but also of any particular apartment on a sheet of paper. However, as a plane has but two dimensions, viz. length and breadth, and a solid three; they cannot all be represented by only one projection on the same plane.

A plane is an even surface, to which a right line may be every way applied, and upon which

there are several ways of projecting solids. We shall only treat of the orthographic projection, as best suited to our purpose.

Before any solid can be represented by this way of projection upon a plane, it must be supposed to be cut by several planes...”

Mungo Murray describes the orthographic projection system and its properties, solving diverse problems. The accuracy level of the drawings is quite good (figure 5), and complex geometry and representation knowledge are applied for the hull lines drawing.

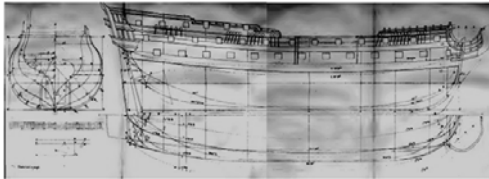


Figure 5: Plates from Mungo Murray treatise.

Architectura Navalis Mercatoria, (1768) [8] has remarkable collection of sophisticated line drawings (figure 6) that offers a fascinating treatise, documenting merchant and naval ships from various countries, it features 70 illustrations that chart vessel dimensions, crew size, storage capabilities, and manner of rigging.

Through the documentation recovered from the shipbuilders, it can be observed that geometry and most aspects of the orthographic projection system are applied.

It can be appreciated that the main parameters to construct a galleon vase are the curves shown in figure 7 and the ribs. Ribs shape was made with circumference arcs. In 1851, in *A history of naval architecture* [10] it is said that “at Paris, in 1681...M. Renaud formed the curves of ships by means of the conic section. On this subject, M. Bouguer remarks, that these curved lines were not preferred because they were discovered to possess any peculiar advantages which rendered them more suited to this use; they were employed because they

were better known, or easier to describe: they presented themselves first.”

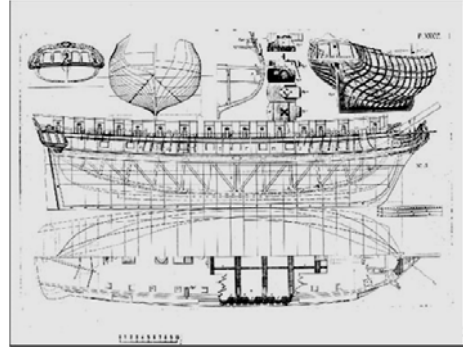


Figure 6: Complete and sophisticated drawings where hull lines plan are as today's.

In figure 8 a sixteenth century galleon is represented with a characteristic drawing of the XVII century, when this book was written.

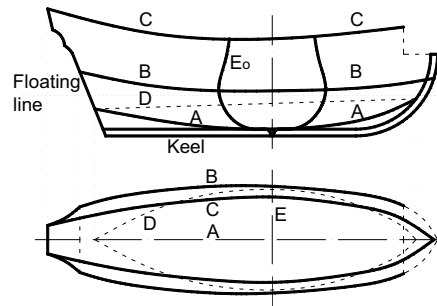


Figure 7: Side view and plan view of the rising-lines plan of a galleon. The guidelines for the curves used in constructing the ribs are as follow: (A) is the lower rising-line, (B) the breadth line, (C) the toptimbers line, (D) the lateral edges of the floor; (E) the midship frame.

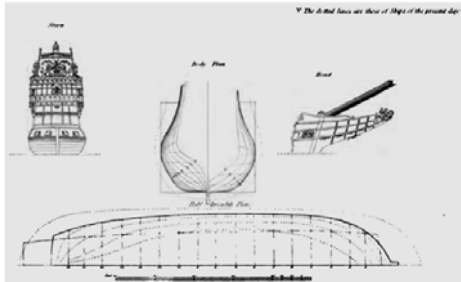


Figure 8: The Royal Sovereign (1591), body and half breadth plans, from A History of Naval Architecture, by John Fincham (1851).

4. MEASURE OF HULL POINTS.

The building process of a ship hull and the result obtained depends on the number of points that are defined and their accuracy level. Similarly, it is relevant in reverse engineering.

In the following paragraphs, two different techniques are explained for the process of a 3D model generation: a total station and a laser scanner (Figure 9).



Figure 9: Total station Leica TCR705. Laser scanner Leica HDS3000.

4.1 Total station measurement.

For the measure of a ship with a total station, a Leica TCR705 machine has been used, which has the capacity of measuring distances without prism. Its technical characteristics are: Compensator accuracy: 1.5", Magnification: 30x, Accuracy standard deviation: 5" y EDM accuracy: 5mm+2ppm.

For the hull modelling with a total station, two different methodologies have been applied:

The first one consists on measuring an irregular group of points by taking three points with local coordinates as a reference (figure 10). With this irregular cloud of points and digital models based on the Delaunay algorithm, a hull model has been built. Then, it is possible to get the profiles according to the ribs. This process is fast and simple, but the result obtained is not worthy enough.

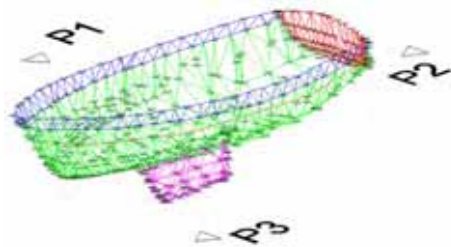
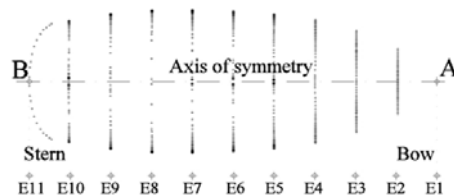


Figure 10: Hull modeling with an irregular cloud of points obtained by a total station.

The second work carried out with the total station consists on measuring individually each of the ribs.

For this purpose (figure 11), the two end points located on the bow and stern of the ship have been projected on the floor, A, B. These two points define a horizontal line of the hull symmetry plane. Then, points E1 and E11 are obtained, which are displaced perpendicularly from the axis. Between them 9 more points are established from E2 to E10, from which is measured each of the ribs.



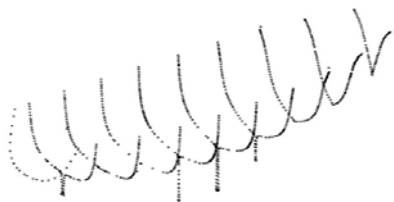


Figure 11: Ribs points obtained by means a total station.

The instrument was settled on each of the points (E1 to E11), and each of the ribs was measured by focusing the instrument in the perpendicular direction of the symmetry axis. These data was captured on the starboard side of the ship, so it was necessary to record the two ends on the bow and stern in order to obtain, by symmetry, the port side.

Other points of the hull are acquired by means of the nurbs curves (the traditional way was manually using templates). Using modelling parameters, the rest of the hull surface between the ribs is calculated. In figure 12 it can be observed the modelled hull using a grade 3 nurbs surface and the control points selection $u=11$ and $v=11$.

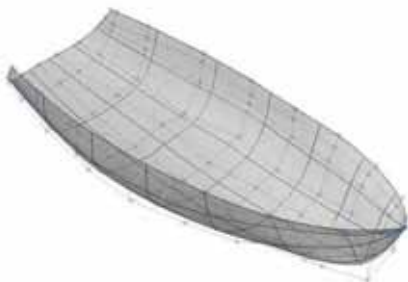


Figure 12: Modeled hull.

The accuracy of the measure depends on the following parameters: magnitude to be measured, measuring instrument and process, person who makes the measure, analysis and error quantification. This is why a detailed study of all of them is required. Figure 13 shows the graphical comparison between a rib measured

with the total station and other obtained by intersecting with the modelled hull. Detail A show where both ribs most differ, having better adjust in other points of the rib. The error that this procedure gives as a result is of the same magnitude of the error of the measuring instrument.

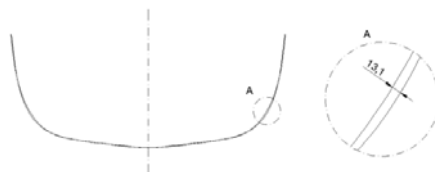


Figure 13: Difference between measured rib and modeled rib.

Once the hull model is calculated, (where keel and rudder are not included, as they are suggested to measure separately), a plane of lines is obtained for the ship appearing in figure 14.

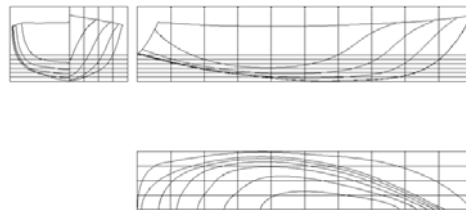


Figure 14: Hull lines plan obtained.

4.2 Laser scanner measurement.

4.2.1 Measurement.

The other 3D modelling technique carried out was the laser scanner Leica HDS3000. Its characteristics are the following: phase scanner with oscillating mirrors and 15 mm resolution at 5 m distance; 6 mm EDM accuracy and 2 mm modeled surfaced precision.

For the scanning process, the instrument (figure 15) has been firstly positioned on one side, B1, surrounded by four fixed reference points, T1 to T4. This allows linking and de-

fining the successive positions where the scanner will be placed. After scanning the port side, the same process is repeated on the starboard side, B2. The scanning resolution was 15 mm, and the measuring time was about the fourth part spent on the total station.

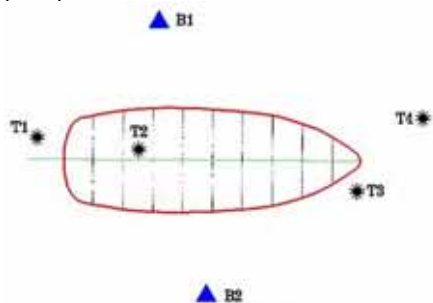


Figure 15:

This kind of instruments is able to build models with a very high number of points. In this case, the model was generated with more than 220,000 points. Based on it, it is possible to get the rib profiles or in any other direction, as it was explained for the previous modelling but obtaining a higher quality and accuracy.

Figure 16 shows the cloud of points obtained. Some “shadow” areas can be observed, which are the areas where the ship was supported.

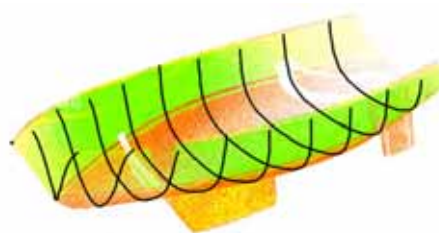


Figure 16: Cloud of points obtained by laser scanner. Ribs obtained from the cloud of points.

4.2.2 Information management. Other applications.

Data capturing is currently being analyzed in order to improve it and solve the low density of points in some areas, due to the angle of incidence. The information that each point registers when scanning corresponds to the dimensions or measure and the reflectance value. The management of the cloud of points requires specific and powerful equipment in order to take advantage of the potential of the information obtained. The big quantity of points allows generating the surface model and improves the dimensional accuracy that the points individually have.

Reflectance depends on certain surface properties and on the angle of incidence. Colors displayed in figure 16 are related with reflectance, and they are not the real colors of the

This information permits extracting additional information referring to the type of materials, surfaces, constructive defects, ship degradation and others.



Figure 17: Measured ship.

5. CONCLUSIONS

The objective proposed in this paper consisted on explaining what a hull lines plans are at the present time, what has been the process followed historically to achieve this kind of hull representation and finally show new methods for 3D model construction, data verification and monitoring during construction.

Multiview orthographic projection was well developed by ship constructors many years before Gaspard Monge wrote his famous work.

The measurement method by means of total station is a good instrument for modeling the hull. It is required a similar time that used manually, but its accuracy is higher and applying the appropriate software can make and modified the modeling of the hull.

The measurement with Laser scanner, requires a more expensive equipment, but is faster, has higher accuracy and further the modeling of the hull, it is expected to perform other applications such as detection of construct defects or aging of the hull

ACKNOWLEDGMENTS

Authors would like acknowledge to our fellows of Cartographic Engineering Area Phd Manuel de Luis Ruiz, Phd Felipe Piña, Phd Candidate Ruben Perez Alvarez.

REFERENCES

[1] F. Oliveira. Manuscript: O liuro da fábrica das naus. *Portugal national library*, 1580.

- [2] M. Baker. Manuscript: Fragments of ancient English shipwrihty (Pepysian Library). <http://www.mhs.ox.ac.uk/staff/saj/thesis/baker.htm> [Accessed June 3, 2014], 1586.
- [3] P. Kirsch. *The Galleon. Conway Maritime Press Ltd*, London, 1990
- [4] Deane's doctrine of naval architecture, 1670. *Conway Maritime Press Ltd*, London, 1981.
- [5] M. Bouguer. *Traité du navire, de sa construction et de ses mouvemens*. Quay des Augustins, Paris, 1746.
- [6] L. D. Ferreiro. *Ships and Science: The Birth of Naval Architecture in the Scientific Revolution, 1600-1800*. The MIT Press Cambridge, Massachussets, USA, 2007
- [7] Mungo Murray. *A treatise on ship-building and navigation*. Printed by D. Henry and R. Cave, London, 1754.
- [8] Fredrik Henrik Af Chapman. *Architectura Navalis Mercatoria*. <http://covax.bth.se/chapmanprojekt/ritningar.html?> [Accessed June 3, 2014], Sweden, 1768.
- [9] Marmaduke Stalkartt. *Naval architecture*. Printed by J.Boydell Cheapside, J. Dodsley Pall Mall & J. Sewell Cornhill, London, 1781.
- [10] J. Fincham. *A history of naval architecture*. Published by Whittaker and Co, London, 1851.
- [11] J. Cavada, F. Fadon. Application of laser technology for measurement and verification of marine propellers. In *Proceedings of the ASME 2012 11th biennial Conference on Engineering Systems Design and Analysis (Nantes, France, july 2-4)*, Nantes, 2012.
- [12] F. Perez. Modelado Geométrico del casco de un buque. *Proceedings of the XVI International congress of graphic engineering (Valladolid, Spain, June)*, Valladolid, 2001.
- [13] Lich, D. et al. "Benchmark Tests on a Three Dimensional Laser Scanning System".

Geomatics Research Australasia, N° 72, pp. 1-24, 2000

- [14] Bohm, J.; Schuhmacher, S. "First experience with the Laser Scanner Leica HDS 3000" *Geo-Information-System*, Vol. 18, N°4, pp: 36-44. 2005.

ABOUT THE AUTHORS

1. Fernando FADON, Phd (Eng.), is professor at the University of Cantabria in the Industrial Engineering College. His e-mail and postal address is as follows: [fadonf@unican.es]; Department of Geography Engineering and Graphics Expression Techniques Av. Los Castros s/n 39005 Santander (Cantabria) Spain.
2. E. CERON, Phd (Eng.), teaches at the University of Cantabria in the Industrial Engineering College. His email and postal address are: [ceronje@unican.es]; Department of Geography Engineering and Graphics Expression Techniques Av. Los Castros s/n 39005 Santander (Cantabria) Spain
3. R. PEREDA, Phd (Eng.), is a lecturer professor at the University of Cantabria in the Civil Engineering College. His email and postal address are: [raul.pereda@unican.es]; Department of Geography Engineering and Graphics Expression Techniques Av. Los Castros s/n 39005 Santander (Cantabria) Spain.
4. L. FADON, Forestry Engineer and Master of Science in Geographical Information Management.

HYPERCAL^{3D}: A COMPUTER APPLICATION TO SUPPORT THE TEACHING AND LEARNING OF DESCRIPTIVE GEOMETRY

Fábio Gonçalves TEIXEIRA and Sérgio Leandro DOS SANTOS

Federal University of Rio Grande do Sul, Brazil

ABSTRACT: This work presents the HyperCAL^{3D}, an application to support the teaching and learning process of Descriptive Geometry through the study of solid objects. The methodology used for its implementation and the main features of the application are described. A selection of concepts was carried out to determine the functional structure of the software should have. From this, the main features were modeled through processes of vector geometry equivalent to that used in Descriptive Geometry. The main features introduced include: the projection process, the representation of hidden lines in three-dimensional model and the projections, successive auxiliary views in real time, and the intersection process. All these tools are implemented in an application that aids the learning process of the students and the teaching procedures of professors.

Keywords: HyperCAL3D, Descriptive Geometry, Computer-based Learning.

1. INTRODUCTION

The Descriptive Geometry (DG) is a key discipline to the project activity. Systematized by Gaspard Monge in the late eighteenth century, DG is the basis of orthographic views and modern Technical Drawing (TD) [6]. The development of DG, and therefore of TD, leveraged the industrial revolution, because it made possible the development of machines and equipment with precision mechanisms, which was impossible until then without the graphical methods for geometric design of parts and structures. Because it is so important to the project activity, DG is part of the curricula of Engineering and Architecture courses since practically its creation, and more recently, of the Design courses as well.

Despite its importance, the curricula of these courses are reducing the workload of DG and in many cases even eliminating it. With the advent of CAD (computer-aided design) systems, there are those who consider that there is no need to learn the graphical processes theory, because they are made automatically by computer applications. However, it is an illusion, because the CAD systems do not make the de-

signs for the designer. He is the one who must have spatial understanding to design in 3D. CAD is a valuable tool, but without a skilled user, with knowledge of the graphical processes and refined spatial understanding, it becomes useless. Therefore, eliminating Descriptive Geometry from these courses, will result in inadequate training of designers, because they will not versed in tridimensional space, which should be one of the primary skills for design professionals (Engineers, Architects and Designers).

The reduced workload is also a problem for teaching DG, because it is a discipline with a dense content and requires great abstraction reasoning of the student. Thus, the workload should be sufficient for proper training of all concepts and processes of the DG. In the courses of the Federal University of Rio Grande do Sul (UFRGS), Porto Alegre, the workload is generally 60 hours. However, some courses have only 30 hours. Given this low workload, there is no way to expect a high level of learning, because the traditional approach of DG requires a high degree of abstraction from the start, and there is not enough time for using

contextualized applications of the concepts.

Despite of being nearly 220 years old, there was little innovation in the teaching of DG for over 2 centuries. There are isolated initiatives, but there wasn't still a paradigm shift that revolutionized the teaching of DG. The main initiatives are based on the use of computer technologies through two and three-dimensional virtual models that illustrate the processes of traditional DG [5, 9, 11].

In contrast, some authors have proposed alternative ways of using technology, drifting away from the mere illustrations by searching for new ways to teach DG. An example is the Descriptive Geometry application, which uses an environment where DG problems are solved through a dedicated programming language based on [2] and [4] present initiatives with the use of CAD systems for describing complex processes in DG.

However, it is from the work [12] that emerges the promise of a paradigm shift in the teaching of DG. In this paper, the authors propose a new approach - entirely based on solid objects, a new nomenclature - aligned to the concepts of three-dimensional geometry and a process that seeks to enhance the student's experience with three-dimensional elements and processes in an integral way. The proposed approach seeks the concrete experience before the introduction of abstractions.

With the proposed new approach to the teaching of DG, there was a problem: how to provide a concrete experience and at the same time, teach the key concepts of descriptive geometry? The hypothesis to the solution was that the concrete experience and the education should happen in the most varied forms, such as physical models, virtual models and computer simulations of graphic processes in an interactive three-dimensional environment.

Thus, the main objective of this work is the development of an interactive computer program that provides a concrete experience and also provides an understanding of the fundamental concepts of DG.

The result is HyperCAL^{3D}, an application to

support the learning of DG. This computer application is innovative because it was developed specifically for teaching DG and simulates its processes in a three-dimensional environment, in real time, through direct manipulation of the elements. In addition, the whole application, which is a direct application of Computer Graphics, was developed using the concepts of Descriptive Geometry in a vectorial approach. Therefore, the design of the application itself is totally based on the concepts of DG.

The HyperCAL^{3D} is designed to be used by teachers to create exercises and for content exposure, and also by their students, to problem solving and visualizing processes interactively. This article presents the main concepts used in the development of HyperCAL^{3D}, and show examples of applications of the application for teaching DG.

2. METHODOLOGY

The methodology for the development of HyperCAL^{3D} involved a vector analysis of the DG graphic processes to synthesize their computer simulation. The following describes the main stages of the development process:

- Selection of the concepts of DG to be supported by v;
- Analysis of the DG graphic processes through decomposition in its primary functions;
- Summary of the vector operations equivalent to the basic functions of graphics processes;
- Development of a data structure to allow the representation of graphic objects and the associated vector operations;
- Computational implementation through an interactive graphical interface.

Below are detailed some of the main aspects of these methodological steps.

2.1 Selection of Concepts

The stage of selection of the concepts is related to the curriculum of the DG discipline that in-

cludes traditional concepts and some more advanced concepts usually not related to this subject.

All traditional basic concepts were covered, including projection, projective systems, Multiview projection system. Concepts traditionally known as auxiliary views are called here Change of Reference System (CRS). This change in nomenclature is used to approximate the DG of mathematics and computer graphics that use the same keywords.

Other concepts added are the perspective views and axonometric views as an application of the CRS. This approach allows to relate the concepts of DG with the concepts of perspective projection, showing the source of this knowledge.

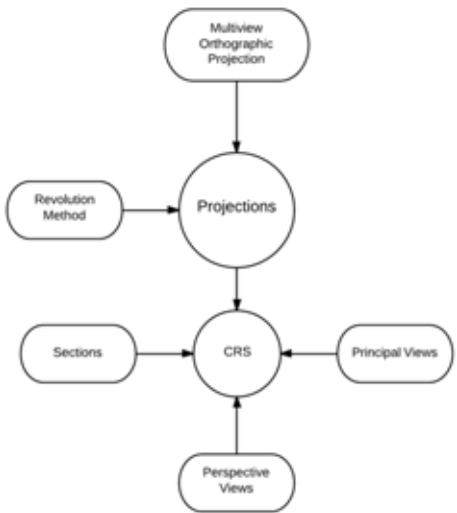


Figure 1: Proposed relationships between the concepts of DG.

Another topic related to the perspective projections is getting the main views of a solid by Successive Changes of Reference System (SCRS). In this case, we start from some perspective views of the solid aiming to obtain the main orthographic views.

The other concepts listed in Table 1 do not

present novelties compared to traditional concepts of DG.

2.2 Concept Analysis

Through a careful analysis of the concepts, it was observed that the fundamental concept is the projection; it is the basis for the representation and transformations. Then, the concepts of revolution and multiview system that are directly associated with the projection. On another level is the operational concept Change of Reference System (CRS), which uses the concept of projection to generate transformations. In a last level are the concepts of intersection, cuts and sections, which use the concept of CRS as a base. Figure 1 shows a diagram of the relationships analyzed.

In the next step, these concepts are unfolded in vectorial processes in order to simulate in a computer environment the graphical processes of DG.

2.3 Summary of Main Vector Operations

The main vector operations built from the analysis are aimed at obtaining: projections; Changes of Reference System; Revolving of objects around an axis; Multiview orthographic projection representation; visibility of edges and faces; intersections and cuts.

Virtual Projection System – The process used in HyperCAL^{3D} to make the projections uses the vector geometry to accomplish the same graphical process used in DG. The algorithm developed to perform the projection considers that each projection plan has an associated Local Reference System (LRS), where the XY plan of the LRS is coincident with the projection plan. Thus, each vertex of the object to be projected has its coordinates computed in the LRS of the projection plan. The projection of each vertex is computed considering as zero the Z component of its coordinates in the LRS and then computed again to the Global Reference System (GRS). The projection of the object is generated by joining the projections of the vertices according to the topology of the

solid model. This procedure is extremely simple and efficient for the generation of projections.

The projection system developed for the HyperCAL^{3D} uses the same logic of the Mongean system, where each projection plan is orthogonal to the adjacent plans, forming a dihedral with double projection. This system allows an unlimited number of projection plans, such that the Reference Systems (RS) are always defined by pairs of plans, one with the horizontal projection (HP) and another with the frontal projection (FP) of the object (Fig. 2). Each projection plan has created a structure of classes and methods that reproduce the vector operations described for performing the projections.

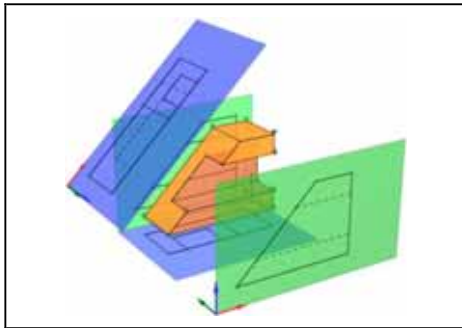


Figure 2: Examples of projection plans and their reference systems.

Change of Reference System (CRS) – which is traditionally known as auxiliary views, uses the same projection process described in the previous item. In this case, this procedure is the vectorial equivalent to the graphic process that trace perpendicular lines to the folding line and get the coordinates from the previous projection plan. Thus, The Projection System developed for the HyperCAL^{3D} already holds natively the CRS process, since it allows multiple simultaneous projection plans. The big innovation here is the interface which allows the user to enter a new projection plan directly and manipulate its position in real time interactively

with the mouse.

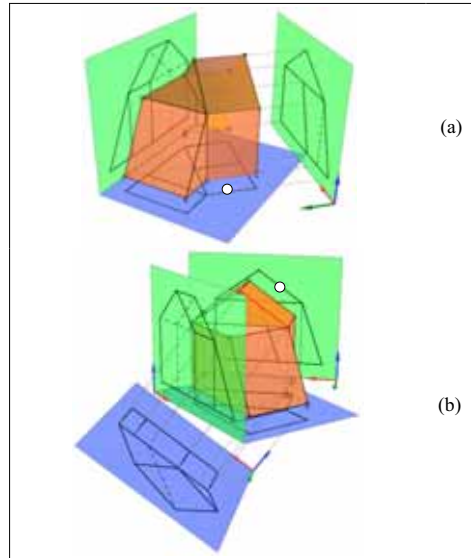


Figure 3: Successive auxiliary views: (a) 1st plan perpendicular to a selected direction. b) Second plan perpendicular to the previous and parallel to the edge view of the face.

Representation of Hidden Lines – the correct representation of solid objects, implies in the adequate representation of hidden lines, those that are not visible by the observer. In HyperCAL^{3D}, visibility is analyzed in the 3D model and its projections. The solution adopted uses similar approaches in both situations and, most importantly, uses the same technique of DG to verify the visibility of lines in orthographic projection.

To check the visibility of a line on a projection plan in DG the intersections of lines are analyzed, and at the points of intersection of the projection lines, it is verified which line is farther from the projection plan (or nearest the viewer). This is the visible line and the other line is, therefore, invisible. This process is widely used in teaching DG and solving visibility problems on the projections of objects. In HyperCAL^{3D}, a similar technique was adopted.

The visibility of the solid model uses an al-

gorithm of ray casting type [14], which traces a line from any object to the camera and checks the intersection of this with the objects of the virtual scene. So, you can check if there are faces or edges between the point of origin of the line and the camera and the distances between them. If present, the point is not visible. Otherwise, the point is visible. Thus, for each edge of the model a sample of points is taken and checked for visibility. May be three cases of visibility: totally visible, totally invisible and partially visible. In the first two cases, the processing is concluded with verification of sample points and the visibility of the edge defines the line type to be used in the representation, continuous line, for visible, and dashed to invisible.

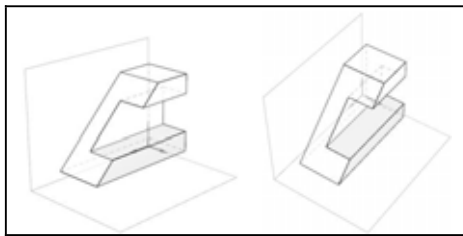


Figure 4: Example of edges with total, partial and completely invisible visibility. Source: the author.

If the edge is partially visible, it is necessary to determine the point of change in visibility (PCV), to properly define the portion with a dotted line and the one with a full line. The process used to find the PCV in the edge is an edge bisection, iteratively repeating the procedure described above checking the visibility of the edge. The length between the two points whose visibilities are distinct is divided in half and the visibilities for the endpoints of the two resulting segments are checked. The segment that has endpoints with different visibilities is selected to restart the process of bisection. This process is repeated until the remaining segment reaches a dimension smaller than certain precision limit. Figure 4 shows the representation of

the hidden lines of a model from different viewpoints.

Multiview Projection – To build multiview projection (MP) automatically is an extremely useful teaching tool in DG, as it allows the visualization of conventional representation process in two ways, in 3D and 2D. The HyperCAL^{3D} performs this procedure directly from the three dimensional model with all plans entered by the user, simulating exactly the situation of generation of MP in DG (Fig. 5).

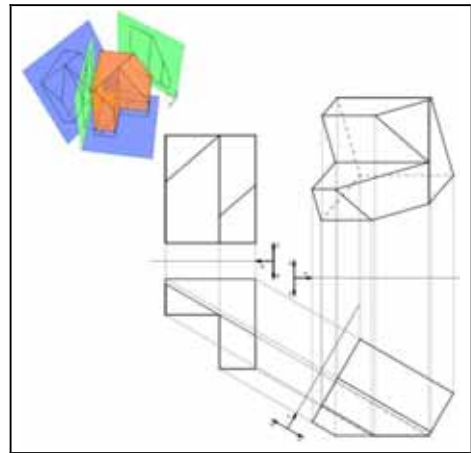


Figure 5: Example of generating MP from the 3D scene.

The MP of the objects is made folding the projection plans one by one, beginning with the last inserted plans. Each plan is folded over its predecessor plan. The folding of the plan of projection turns on an angle of 90° around the folding line, including all projections contained therein. This process of folding is a vector procedure, where is used a rotation matrix [8] coupled to the folding line (x-axis) of each RS. To be able to perform all the folding robustly, the projection plans are organized in a hierarchical data structure of tree type, where the root is the horizontal plan. All other plans are branches that converge to the root.

After assembling the tree, each plan is fold-

ed over the predecessor plan, on the tree, starting with the outermost branches, finishing in the horizontal plan, the only plan that is not folded. For a plan to be folded down, all his descendants should have already been folded. Another important rule is that when a plan is folded down, all their descendants are too. This succession of folding organized and targeted by this tree-like structure results in the complete flattening of the projection system.

2.4 Data Structure

All features of HyperCAL^{3D} rely on a data structure that supports both the representation of the objects as the operations on them. The data structure developed was defined as a hierarchical structure of classes, which organizes all the objects and their operations.

The basic geometric classes are Vertex, Edge, Face and Solid. Whereas the edge has two vertices as class properties. The Face class has a list of vertices and edges as properties. The Solid class has a list of vertices, a list of faces and a list of edges as properties. The Solid class has methods that guarantee the data organization of Vertices, Edges and Faces. Furthermore, the solid class has methods for determining the visibility in the 3D environment.

The “Projection System” consists of a class that contains a list of objects of the class “Projection Plan”. This class, in turn, contains the methods to project the geometry of the geometric objects, creating objects of class “Projection”, which become properties of the Projection Plan. Moreover, this class has methods that compute the visibility of projection lines and to fold the plan and all the projections contained therein. The projection system has methods to control its contained plans, including insertion, removal and folding of plans, to generate the multiview projection.

2.5 Graphical Environment

The programming language chosen for the computational implementation is the Embarcadero Delphi®, which enables object orienta-

tion and has a compatible performance in interactive applications. The graphics processing technology chosen is the OpenGL® [15], which is a standard graphics library in the computer graphics industry.

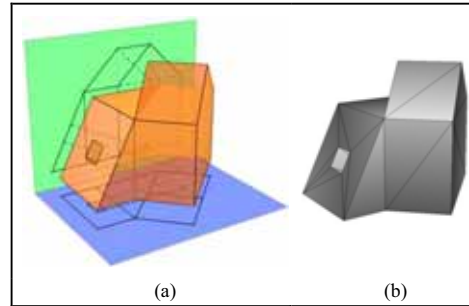


Figure 6: Representation of solids in HyperCAL^{3D}. a) Form presented to the user. b) Faces represented by the mesh on the data structure.

The data structure for the representation of three-dimensional models is based on a hierarchical representation of objects. Thus, a solid is divided into faces and these are subdivided into triangles mesh. This is necessary because there is geometric features in not convex solids, as the holes and indentations that can derail representation by a single polygon. Thus, each side of a solid is actually a mesh of triangles, which enables the representation of faces with complex geometry.

The triangle mesh is constructed by a mesh generator specially developed based on the technique of advanced front [10] and Delaunay triangulation [1]. Figure 6 shows the representation of a solid in the form that is presented to the user (Fig. 6a) and the representation of the triangle mesh (Fig. 6b), which is not perceived by the user.

The three-dimensional environment built allows large interactivity. You can move the camera dynamically, select objects for various purposes in the process of interaction with the user, such as inserting projection plans, for example.

3. RESULTS AND APPLICATIONS

The result of the development process is an application with features that let you work with various concepts of Descriptive Geometry (DG), enabling teachers to expose the contents directly in the program interface.

There are several applications in DG that can use the HyperCAL^{3D} as a support tool:

- visualization of 3D models with the MP for understanding the geometry;
- determining of true length projection (TL) of oblique segments by primary CRS;
- determining true shape (TS) of face (plan) with edge view projection by primary CRS;
- determining of the edge view projection of an oblique plan by primary CRS;
- determining TS of an oblique plan by double and successive CRS;
- determining of distances between a point and a line between lines, between a point and a plan, and between plans;
- determining axonometric view from any point of view;
- determining the principal orthographic views of a solid from successive CRS.

Following are described: the interface of the application and some significant applications of the teaching of DG that can be used with the support of HyperCAL^{3D}.

3.1 The HyperCAL^{3D} Interface

The graphical user interface HyperCAL^{3D} presents three main areas: graphic area, control panel and toolbar. The toolbar is the interface access to the main commands.

The side panel is a direct interface with many features. This panel contains multiple pages, accessed by the tabs at the top. Where it is possible to insert vertices (points), connectivity of the faces (Faces), position the camera precisely (Camera), configure the interface elements (Setup), verify characteristics of the model (Model) and find the intersection between the solid and plan (Intersection).

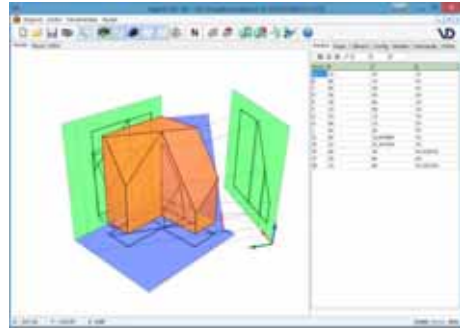


Figure 7: Graphic user interface: Model.

In the graphical area, which is dominant in the interface, there is a panel with two pages. On the Model tab, the 3D interface, which is the main interface, is presented. In this area, the solid is displayed in 3D along with the plans of the Projection System. That is where the interactive camera positioning commands, object selection, insertion of projection plans for Change of Reference System (CRS) are performed.

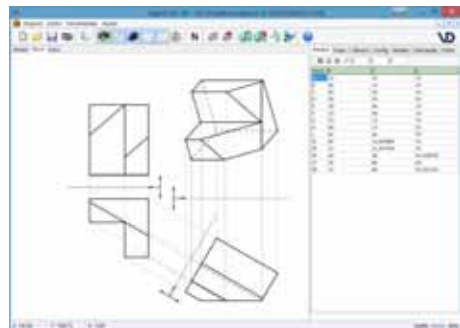


Figure 8: Graphic user interface: MP.

Objects have specific colors, according to their type, but you can change the representation for black and white. In MP tab (Fig. 8), the current scene corresponding to the unfolded system is presented, where projection plans are unfolded until all projections are on the same plan, just as in the DG drawings. Both, in the

model area as in the MP, commands can be displayed, such as Interactive Zoom and Pan, using the scroll button of the mouse. You can switch between the Model and MP directly and quickly, allowing the comparison between the two situations thus engaging students thinking.

3.2 True Length of Oblique Segments

HyperCAL^{3D} has no specific commands to a particular situation and do not automatically performs all tasks for the user. Thus, determination of True Length (TL) of an oblique line is conducted through an appropriate CRS, which must be set by the user/student. So it is still necessary to understand that the plan to be inserted into the CRS must be parallel to one of the projections of the line segment.

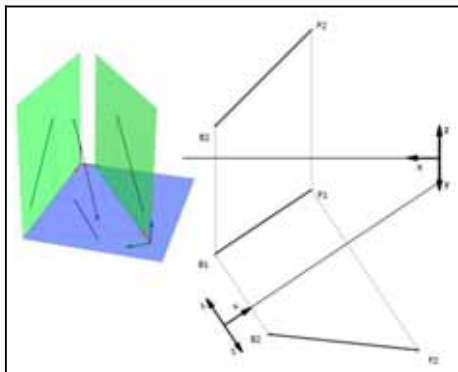


Figure 9: TL of an oblique line in 3D and MP.

Likewise, it is the users/students decision to make a CRS to obtain the true angle in the Horizontal Projection Plan (HP) or Frontal Projection Plan (FP). Therefore, it is not the application solving exercises for students. The application is a supporting tool, but without the knowledge of DG, the problems cannot be solved, with or without the application. Figure 9 shows an example of obtaining the FSP of an oblique edge (GC) of a solid. In this case, there was a SCRS through a new FP parallel to the horizontal projection of the edge. One can see the process in 3D with the entire solid, where a

FSP of the line it is invisible on the auxiliary view. In MP, the option of isolating the edge was used, leaving visible only the edge under consideration.

3.3 Determination of True Shape of a Face with an Edge View Projection

To determine a True Shape (TS) of a face with the edge view projection [3] is a problem which can be solved with a single CRS, provided it is chosen correctly. There is no specific command for it in the application. The user/student should know that a plan is projected in TS when their edge view projection is parallel to the folding line (FL - X axis). And yet, the student must understand what it is and how to identify the edge view of a face.

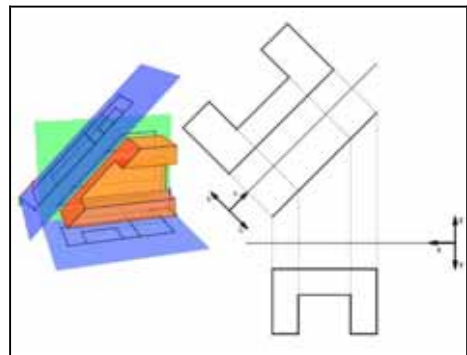


Figure 10: True Shape of an edge view of a face.

3.4 Finding the Edge View of a Plan

The edge view of an oblique face has several applications in DG. The finding of the edge view of an oblique plan is one of the most important concepts for the solution of problems involving plans in DG, which supports the determination of the angle between a face and a projection plan. The edge view is also important whether as an intermediate step for the determination of TS of an oblique plan, or even in situations where it is necessary to determine distances between the plan and other objects.

The HyperCAL^{3D} does not have a specific

command to determine the edge view of a plan. It is the user/ student must make the appropriate auxiliary view to achieve the result sought. Thus, to determine the edge view of a plan, it is necessary to obtain the point view of a line of the plan [3].

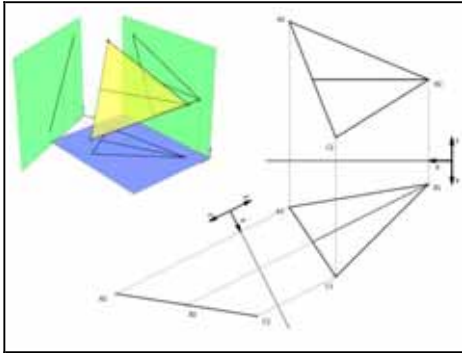


Figure 11: Determining of the edge view of an oblique face in 3D and in MP.

If a line has a point view projection, the adjacent view has a TL projection and perpendicular to FL. To get the edge view of a plan, it is necessary to identify or represent a line in TL on it and then perform the procedure described. Figure 11 shows an example of obtaining the edge view of an oblique plan. In case, a triangular face, where we drew a horizontal line contained in the plan. Then, there was a plan CRS inserting a projection of TL perpendicular to the horizontal line. With this, the line is projected as view point and the plan is projected as edge view the auxiliary plan.

3.5 Determination of the True Shape of an Oblique Face

To determine the True Shape (TS) of an oblique face is one of the main problems of DG. Therefore reduced from projections of the face, you can get their real dimensions only graphical procedures. When a plan is projected in TS means that it is parallel to the plan of projection. This is indicated by the projection be an edge view and parallel to the FL. Thus, two succes-

sive CRS to find the VG projection of an oblique plan are required. The first obtain the edge view of the plan and the second uses a FL parallel to the edge view to get the TS.

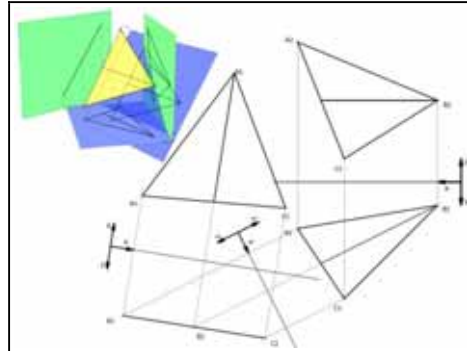


Figure 12: Determination of the true shape of an oblique face in 3D and in MP.

In the HyperCAL^{3D}, the second CRS is made directly selecting as a base plan the projection plan of the first CRS, and as for the direction, selecting the edge view projection of the oblique face obtained in that projection plan. Figure 12 shows the same example of Figure 11, where the edge view of a face was made by a CRS, and a new projection plan was added parallel to the edge view of the face, projecting in this new projection plan the true area of the face. One can observe in MP (Fig. 12) the first Folding Line, used to obtain the edge view and the second parallel to the edge view to get the true shape projection of the face. This example, very simple, using only one face was presented just to illustrate the process. However, it is possible to work with complex solids with the same agility.

3.6 Axonometric Views

Axonometric views are those obtained from orthogonal cylindrical projections, in a way that the three main axes of the object are not projected cumulative in the projection plans. As the orthogonal cylindrical projection, system is the base for DG, it's possible to obtain

an axonometric view in orthographic projection by choosing the appropriate direction of view (line of sight). This line of sight (LS) can be displayed as a vector from the observer to the object, as the direction is important as well, defining from where and to where one is looking at. If the observed object has its main axes parallel to the axes of the projection system, the LS must be oblique in order to get an axonometric view.

The axonometric view appears in the plan that the LS has view point projection, implying that the resulting projection is defined by the direction of the LS. Thus, once one has the direction of observation, it is necessary to make two successive CRS in order to obtain the point view of LS, getting the axonometric view as a result. Figure 13 shows an example of determining the axonometric view in the HyperCAL^{3D} from a given LS, by two points.

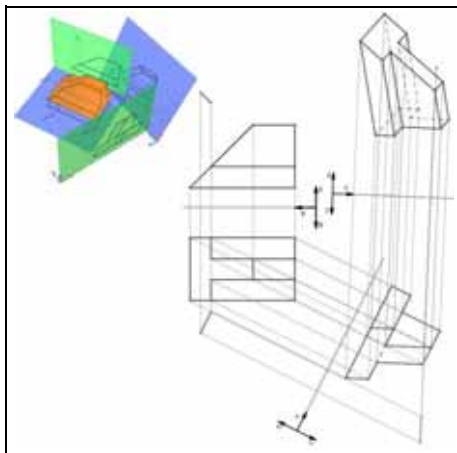


Figure 13: Axonometric view from the line of sight (LS) in 3D and in MP.

The HyperCAL^{3D} interface also has the possibility to get the same axonometric view built by SCRS inside the 3D environment through a proper positioning of the camera. Thus, if the camera is positioned with the same parameters of the LS, the window shows the 3D solid

model with the equivalent axonometric view. This is not a MP but a Computer Graphics tool that can be used to check the result done through DG.

3.7 Principal Views

The principal views of an object are projections aligned with the axes of Reference System and positioned according to specific criteria recommended by the literature and by the rules of Technical Drawing. The principal views are important because they represent objects in views where there are true projections of its dimensions, allowing its physical construction and allowing interventions or changes to its geometry in the design stage. The principal view, in general, is the front view, which characterizes the position of use, or presents more detail and fewer invisible lines. The other views are determined based on the main view.

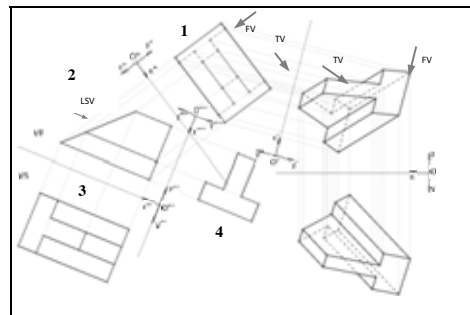


Figure 14: Principal views.

When an object is represented on projections that do not match its principal views, these can be obtained from Successive changes of Reference System (SCRS), until the principal views of the object is obtained. The process of getting the principal views by SCRS starts by choosing the directions that represent the principal views of the object and what position they have in the RS. This choice is directly related to the selection of views. Thus, by setting the front view (FV), it is possible to understand in what position determined edges and plans

should be. In the example of Figure 14, after the Front View and Top view were chosen, it is possible to understand that the edges that are parallel to the FV, which are oblique, should be positioned as view point after positioning the auxiliary views. Furthermore, the face ABCD must be horizontal when viewed chosen were designed. The first auxiliary view is performed in order to find the TL of FV direction (Fig.14-1).

The second successive auxiliary view is performed to get the view point of FV (Fig. 14-2). This is the first principal view. You still need to get the top view of the part, so that the piece is fully aligned with the SR. Obtaining the top view (TV) is made through a third successive auxiliary view, creating a new line perpendicular to TV direction in the second auxiliary view. Thus, the bottom face is positioned horizontally as in the third SR, which is compatible with the chosen top view (Fig. 14-3).

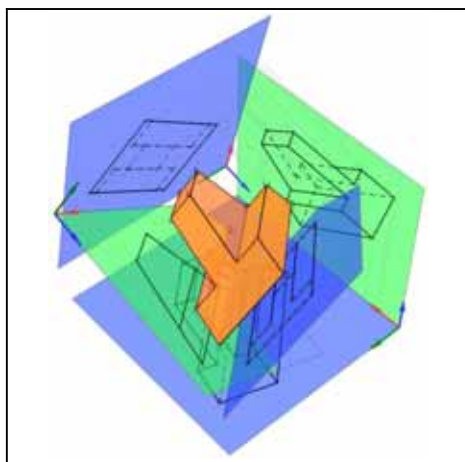


Figure 15: Finding the Principal views in 3D.

From these views, you can recognize the shape and measure height, width and depth of the object directly on the projections. A fourth auxiliary view is made to get the left side view (Fig. 14-4). This whole process is performed in

an interactive way in 3D on HyperCAL^{3D} (Fig. 15), where the user/student can test and understand each CRS in real time.

4. CONCLUSIONS

The HyperCAL^{3D} is a computer program based on a set of algorithms based on vector geometry implemented in a three-dimensional interactive graphical environment, which reproduce the methodology and graphical procedures used in Descriptive Geometry. The result is a powerful, unmatched nationally and internationally for teaching this important for vision and spatial skills, critical reasoning in activities related to the design of artifacts in various scales discipline.

The main processes used in the development of the application, which allows viewing, manipulating solid objects and projection plans, actions that enable the use of HyperCAL^{3D} both by teachers and students in their different contexts of teaching and learning have been described. Teachers can use the application to prepare and test exercises, preparing educational materials printed and presentation and also to present content in an interactive way for students. Students can use the application to observe the processes in 3D and relate them to Multiview Projections trying to understand the logic behind this course, which has been taught for more than two centuries based on unquestioned axioms and laws. The observation of processes in 3D and in Multiview Projections, according to the user/student intention itself, enables a critical analysis and reflection on their part, which meets the construction of knowledge itself.

The actual application version is 6.0 but the HyperCAL^{3D} is still in development and new features are being implemented and tested, seeking the quality of teaching Descriptive Geometry.

REFERENCES

- [1] M. K. AGOSTON. Computer Graphics and Geometric Modelling: Implementation & Algorithms. Springer, 2005.

- [2] D.M. CHEN. Application of 3D CAD for Basic Geometric Elements in Descriptive Geometry. *Engineering Design Graphics Journal*, Volume 64: 01, pp. 10 - 17, 2000.
- [3] M.C. HAWK. *Theory and Problems of Descriptive Geometry*. New York: Schaum Publishing Co., 1962
- [4] F.M.C. Jr. The Need(?) for Descriptive Geometry in a World of 3D Modeling. *Engineering Design Graphics Journal*, Volume 62: 03, pp. 4 - 8, 1998.
- [5] Á.J. RODRIGUES, C.J. HAGUENAUER, & G.G. CUNHA. *A Realidade Aumentada no Ensino da Geometria Descritiva*. Curitiba, s.n., 2007.
- [6] E. ROVIDA. *Machines and Signs - A History of the Drawing of Machines*. Amsterdam: Springer Netherlands, 2013.
- [7] E.T. SANTOS, L.N.D. BARROS, & V.C.P.N. VALENTE. *Projetando uma Ontologia de Geometria Descritiva*. Curitiba, s.n., pp. 918 - 928, 2007.
- [8] P. SCHNEIDER & D.H. EBERLY. *Geometric Tools for Computer Graphics*. s.l.: Morgan Kaufmann, 2003.
- [9] R.D. SEABRA & E.T. SANTOS. Developing the Spatial Visualization Ability with a Virtual Reality Tool for Teaching Descriptive Geometry: a Brazilian Experience. *Journal for Geometry and Graphics*, Volume 17, pp. 101 - 117, 2013.
- [10] F.G. TEIXEIRA. *Modelamento Paramétrico e Geração de Malha em Superfícies para Aplicações em Engenharia*. PhD Thesis. Porto Alegre: 2003.
- [11] F.G. TEIXEIRA, R.P.D. SILVA & TLK SILVA. A Learning Environment for Teaching of Descriptive Geometry. In: *Proceedings of International Conference on Engineering Education*. Ostrava - Prague, s.n., 1999.
- [12] F.G. TEIXEIRA, R.P. SILVA, TLK SILVA. & A.T. HOFFMAN. The Descriptive Geometry Education through the Design-Based Learning. In: *Proceedings of International Conference on Geometry and Graphics*. Salvador, s.n., 2006
- [13] P.R. WALKER. *The Feud That Sparked the Renaissance: How Brunelleschi and Ghiberti Changed the Art World*. s.l.:Harper Perennial, 2003.
- [14] A. WATT. *Computer Graphics*. s.l.:s.n. 1993.
- [15] R.S. WRIGHT, B. LIPCHAK, & N. HAEMEL. *OpenGL SuperBible: Comprehensive Tutorial and Reference*. 4th Edition ed. s.l.:Addison-Wesley, 2007.

ABOUT THE AUTHORS

1. Fábio Gonçalves Teixeira is Ph.D at Mechanical Engineering with emphasis in computer graphics and geometric modeling by the Federal University of Rio Grande do Sul (UFRGS) (2003). He is Director of Graduate Program in Design of UFRGS since 2011. He works with research projects related to Virtual Design of products, including computer graphics, numerical simulation of structures, virtual reality and scientific visualization. He operates in research projects for development of new educational technologies covering the following subjects: virtual learning environments, learning objects, descriptive geometry, computer graphics, hypermedia and virtual reality. He was postdoctoral researcher between 2009 and 2010 at Faculty of Engineering of University of Porto in Portugal in the design of adhesive joints and its application in product development.

2. Sérgio Leandro dos Santos is Director of undergraduate courses in Product Design and Visual Design of Federal University of Rio Grande do Sul (UFRGS) since 2013. Assistant Professor, Department of Design and Graphic Expression at UFRGS. PhD Student (Current) and graduated in Design PGDesign-UFRGS (2009). Professor of Computer Graphics at the Undergraduate course in design - UFRGS.

IMAGE PROCESSING EDUCATION IN THE DEPARTMENT OF INFORMATION ENGINEERING, COLLEGE OF INDUSTRIAL TECHNOLOGY - UTILIZATION OF SPREADSHEET SOFTWARE AND WEB SYSTEM

Masayuki MORI¹, Hiroki TOMINAGA², Naoki IIDA² and Kenichi HIROSE²

¹ Osaka University, Japan ² College of Industrial Technology, Japan

ABSTRACT: This paper describes the learning materials and learning environment construction that were developed for image processing education in the Department of Information Engineering at the College of Industrial Technology. As “Experience of the image processing”, we developed learning materials using familiar spreadsheet software to students. Moreover, as “Grasp of the understanding”, we carried out submitting exercises by LMS (Learning Management System). In addition, as “Understanding of terms”, we developed the image information processing web dictionary form slides of lecture and textbooks to provide environment to search terms that students aren't understand. The results of a questionnaire confirmed that the use of these materials in subjects was effective for enhancing learning.

Keywords: Image processing education, e-Learning, Spreadsheet software, Web dictionary.

1. INTRODUCTION

It is necessary for the information technology engineers to learn media information processing by a device with high quality image display and the diffusion of image contents. The main purpose of the education is “Training of the programmer” in the Department of Information Engineering at the College of Industrial Technology. And to achieve a purpose we provide subjects of information processing technology for students in the curriculum. Above all, there is a strong need to media information processing. Therefore, we provide various subjects to students for two years as an important technique. If possible, students should take a subject for media information processing after they learn a basic knowledge of programming skills for “Training of the programmer”. However, because students also have to take subjects of programming skills in the same period in the limited period of two years in the junior college, they will take a media information processing as unripe programming skills. Consequently, it is nec-

essary to support an experience and an understanding of media information processing technology even if programming skills and knowledge of students is not enough.

Therefore, In Department of Information Engineering, we defined that “Experience of the image processing”, “Grasp of the understanding”, “Understanding of terms” are essential elements of education for media information processing, especially image processing. Moreover, we performed system development and learning environment construction to support those. In this paper, we present practice contents by subject about the image information processing with description of characteristics of developed systems and learning materials.

2. IMAGE PROCESSING EDUCATION AT THE DEPARTMENT OF INFORMATION ENGINEERING

The main subjects of image processing education are “Basic experiments of information processing” and “Image processing.” The following are the main topics that are taught in these subjects.

1) Basic experiments of information processing

This subject was given in the second semester of the first year. It involved not only instructions for using software to process images, but also included practical training on image processing that emphasizes numerical expression while displaying visual data in order for students to learn about the basic concepts of information expression and processing.

2) Image processing

This subject was given in the second semester of the second year. In it, students were taught about the mechanism of image processing at the pixel level, which is the smallest component of digital images, as a fundamental of image processing. It was also designed to help students learn about methods for converting images, such as enhancing image quality and extracting edges. There was also instruction about techniques for processing digital images for applications in various fields such as image measurement and image recognition.

3. DEVELOPMENT OF LEARNING MATERIALS AND LEARNING ENVIRONMENT CONSTRUCTION

In Subjects so far, practical training was conducted using various types of software, and there

have also been examinations of instruction methods as well as academic research reports about the educational effects. However, the practical training using the software was only geared toward an interest in presenting images that had been derived from processing; unfortunately, it was not directly linked to providing students with an understanding of image processing.

Therefore, it was thought that learning materials were needed that could enable “Experience of the image processing” which would strengthen awareness of image processing techniques.

3.1 Learning materials for experience-based image processing

Learning materials for experience-based image processing were developed to enable students to gain “Experience of the image processing”[1]. Figure 1 shows the learning materials that were developed. The materials were Microsoft Excel[2] based applications. Input and output processing of image files was prepared beforehand with macros, and students could gain experience with the basic processing parts of image processing. In addition, by displaying the inputted image files with histograms and other types of graphs, the students could observe the image data.

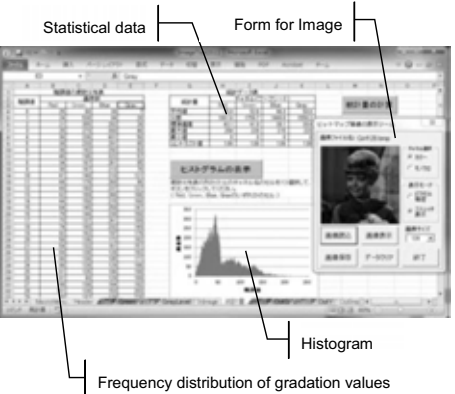


Figure 1: Learning materials for experience-based image processing

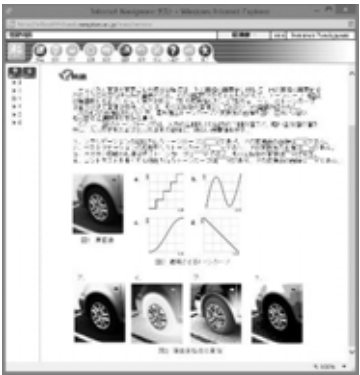


Figure 2: Review problems using LMS

3.2 Review problems using LMS

In the 2005 academic year at this school, LMS was introduced to assist with subjects, and is now used throughout the school. In image processing education, students were asked to solve review problems from the previous lessons that were uploaded into LMS and that they answered at the beginning of subjects to help them “Grasp of the understanding” as taught in the experience-based image processing learning materials. After giving their answers, students could immediately see whether they were wrong or right, which could help them know about their own level of understanding. Figure 2 shows the review problems using LMS.

3.3 Web Dictionary for image processing terminology

The learning materials for experience-based image processing make it possible for students to actually experience image processing, but they do not include detailed explanations about the techniques involved. Therefore, we created a database of technical terms used in image processing, and developed a search system that can enable free browsing. Figure 3 shows the Web Dictionary that was developed. This system is the Web Dictionary, which is run on a browser[3]. The entry words here were terms

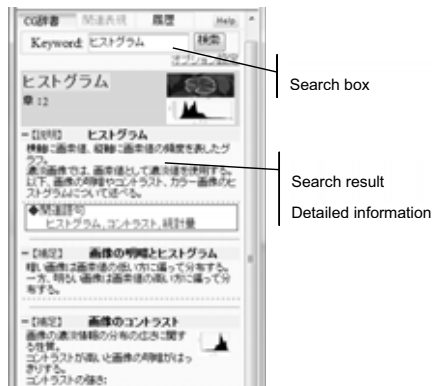


Figure 3: Web Dictionary for image processing terminology

that were used in the “Basic experiments of information processing” and “Image processing” subjects, and detailed information could be browsed through searches. Searching methods included not only entering a key word, but also selecting letter strings from web browsers, Adobe Reader, etc., that could be used to conduct automatic searches. It is thought that this search system can help students to “Understanding of terms”.

4. EXPERIMENT

The experience-based image processing learning materials, review problems using LMS, and the Web Dictionary were used in subject, and the learning results were examined. The following are the results of questionnaires that were executed after each subject.

4.1 Basic experiments of information processing

Because the purpose of this subject was to give students “Experience of the image processing”, a survey was conducted to determine how easy it was to use only the experience-based image processing learning materials. The survey covered 4 years, from 2010 to 2013. Figure 4 shows the results of questionnaires that were executed after the course had been completed. Q1 indicates that the learning materials were easy to use, with a positive evaluation of more than 88% in each year. In the free answer space, many students wrote both that “The materials were easy to use and operation was simple” or that “Operation was intuitively easy to understand.” This indicates that the materials were easy for students to approach. Furthermore, Q2 shows that more than 88% of the responses were positive about the ease of understanding learning materials. These findings indicate that giving students actual experience with image processing is effective for learning.

4.2 Image processing

In this subject, experience-based image processing learning materials and Adobe Photoshop were all used in the exercises and the results were compared for three academic years: 2011,

2012 and 2013. The contents “Image processing” were more technical than those for “Basic experiments for information processing.” As a result, to increase the amount of learning, LMS was used at the beginning of subjects to do review problems which allowed students to measure their own level of understanding. LMS was also used to support classroom tasks such as checking attendance and submitting reports.

4.2.1 Questionnaire about the exercises

Figure 5 shows the results of a questionnaire that was executed after the subject was completed. Q3 and Q4 are questions about the exercises using the experience-based image processing learning materials. In all cases, positive opinions accounted for at least 73% of responses, suggesting that the learning materials were effective for learning. However, Q5 shows that there were many students who felt that the explanations in the learning materials were not adequate. This suggests that besides the experience-based image processing materials, learning materials which contain detailed explanations are also needed.

Q6 shows the results of questions that asked about the effectiveness of the experience-based image processing learning materials and Adobe Photoshop[4] in exercises. According to the figure, the use of Adobe Photoshop in the 2012 and

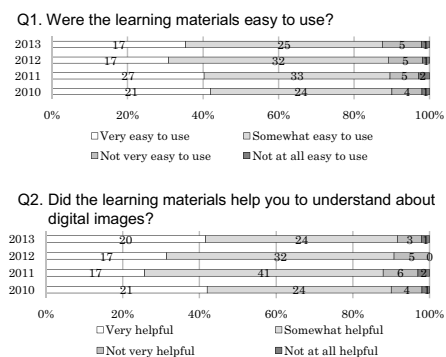


Figure 4: Questionnaire about ease of use

2013 academic years was more effective. Furthermore, Q7 and Q8 show the use of Adobe Photoshop together with the experience-based image processing learning materials was more effective than just the learning materials. In the free answer space, many students thought that the operating methods in the experience-based

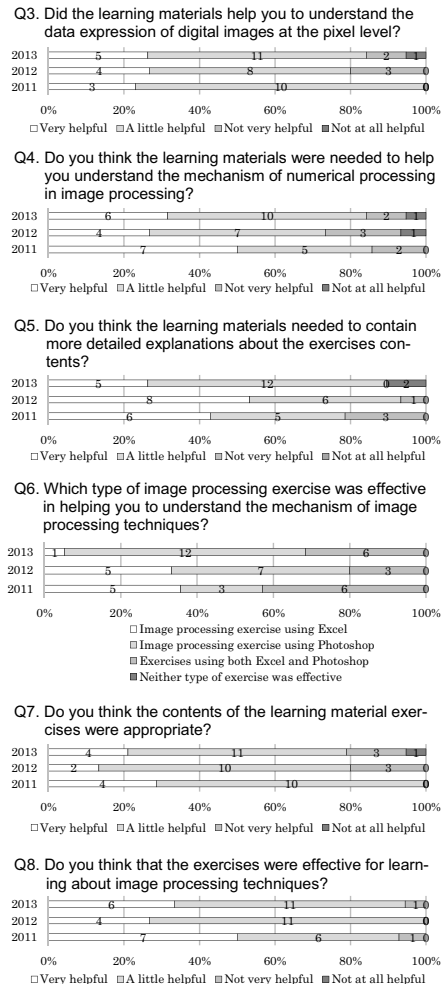


Figure 5: Questionnaire about the exercises

image processing learning materials were difficult, and it appears that Adobe Photoshop, which has higher workability, was easier for them to use. Therefore, it is thought that while the experience-based image processing learning materials had functions for displaying images, when it is necessary to display processing results immediately, it may be necessary to develop lessons that use Adobe Photoshop and the other learning materials in conjunction with the experience-based image processing learning materials.

4.2.2 Questionnaire about LMS

Figure 6 shows the results of a questionnaire that was executed after the subject had ended. Answers were given based on a scale of 1 to 5. Q9 shows questions about the effectiveness of LMS as a learning support. The average for each year ranged from 3.75 to 4.46, indicating that LMS did help to increase the effectiveness of subjects. In addition, Q10 shows that average evaluations

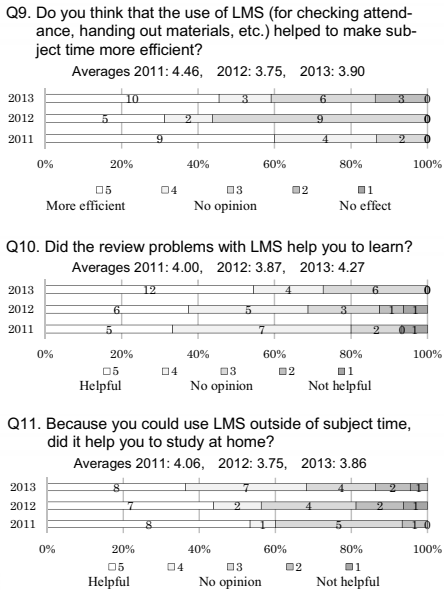


Figure 6: Questionnaire about LMS

of the review problems ranged from 3.87 to 4.27, indicating that they also helped students to learn. Furthermore, from Q11 we can see that the review problems were used for learning outside of subject time. These results suggest that the review problems that were given when subject started not only helped students to understand the materials, but they also helped to increase student's study time.

4.3 Questionnaire about the Web Dictionary

The results of the questionnaire about image processing indicated that, in addition to the learning materials for the experience-based image processing, there was a need for materials that provided more detailed explanations. Consequently, the Web Dictionary was provided to students on an experimental basis as another learning material for "image processing." Figure 7 shows the results of a questionnaire that was executed after the end of the subject in 2011. There were 15 students in the subject. Because the purpose of the Web Dictionary was to provide students with detailed explanations about technical terms used in image processing, comparisons were made with the textbook that was used in the subject. The textbook used in the subjects was "Visual Information Processing[5]" published by Computer Graphic Arts Society (CG-ARTS).

According to Q12, most students responded that they were able to look up necessary information faster with the Web Dictionary. In contrast, for the question about ease of viewing in

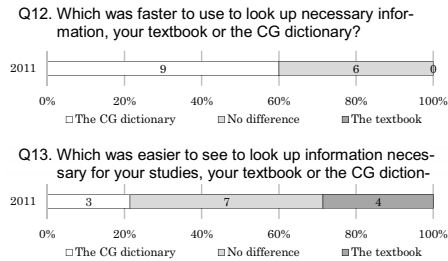


Figure 7: Questionnaire about the Web Dictionary

Q13, most students were supportive of the textbook. Although it is easy to display information with the Web Dictionary's search function, it appears that the display area was so narrow that it made the displayed information more difficult to see, compared with the textbook.

5. CONCLUSIONS

This paper has examined efforts to supplement "Experience of the image processing", "Grasp of the understanding", and "Understanding of terms" in image processing education and has corroborated the results of learning. The results of questionnaires that were executed after the subjects had ended demonstrated the effectiveness of the learning materials we provided and the improved learning methods.

A future topic is to increase study time outside of subject time by promoting the use of LMS, the Web Dictionary, and other aids. Plans call for experimenting with "flipped learning" by focusing on practice with the experience-based image processing learning materials during subject time.

REFERENCES

- [1] Kenichi HIROSE, Naoki IIDA, Toyohisa KANEKO and Kiyotsugu SATOH. Activity Reports on Development of Experience-based Learning Material for Media Processing Education. *Bulletin of College Industrial Technology*. 46: 65-69. 3 2009.
- [2] Microsoft Corporation. Microsoft Excel. <http://office.microsoft.com/en-001/excel/>. accessed 6 2014.
- [3] Adobe Systems Incorporated. Photoshop. <http://www.adobe.com/products/photoshop.html>. accessed 6 2014.
- [4] Tomomi OMAE. Development of User

Friendly Learning Management Systems: WebOCM and Web4u. *Studies in e-learning Language Education*, Volume 3: 1-8, 12 2008.

- [5] Computer Graphic Arts Society. Visual Information Processing. 4 2004.

ABOUT THE AUTHORS

1. Masayuki MORI is an Assistant Professor in the Cybermedia Center, Osaka University. His research interests are Interactive Three-dimensional Computer Graphics and e-Learning. He can be reached by e-mail: morim@cmc.osaka-u.ac.jp, by postal address: 4F 1-32 Machikaneyama / Toyonaka / Osaka / JAPAN.
2. Hiroki TOMINAGA is an Assistant in the Learning Support Center, College of Industrial Technology. His research interests are e-Learning. He can be reached by e-mail: tominaga@sit.sangitan.ac.jp, by Fax: +81-6-6431-7992, by postal address: 1-27-1 Nishikoya / Amagasaki / Hyogo / JAPAN.
3. Naoki IIDA is an Associate Professor in the department of information Engineering, College of Industrial Technology. His research interests are e-Learning. He can be reached by e-mail: iida@cit.sangitan.ac.jp, by FAX: +81-6-6431-7992, by postal address: 1-27-1 Nishikoya / Amagasaki / Hyogo / JAPAN.
4. Kenichi HIROSE is a Professor in the department of Information Engineering, College of Industrial Technology. His research interests are Image Processing, Computer Graphics and Human Interface. He can be reached by e-mail: hirose@cit.sangitan.ac.jp, by FAX: +81-6-6431-7992, by postal address: 1-27-1 Nishikoya / Amagasaki / Hyogo / JAPAN.

IMAGINED SPACE, REAL SPACE. THE CURTAIN OF THE THEATRE AND THE CENTRAL HALL OF THE ROYAL PALACE OF CASERTA

Nicola PISACANE

Seconda Università degli Studi di Napoli, Italy

ABSTRACT: The study proposes a critical reading of the painted canvas curtain behind the proscenium of the Theatre of the Royal Palace in Caserta. The canvas, still not precisely dated by current historiography, depicts an urban scene with a semicircular colonnade ended by a building with a classical porch, as well as a double staircase with handrails and sculptural images at the top. Interesting the point of view for the construction of the accidental perspective that strongly recalls the model theorized by Ferdinando Galli Bibiena in his treatises on architecture and perspective. This spatial system, known throughout Europe as a '*scena per angolo*' (scene for corner), borders the main point of perspective outside the public eye, so over the limits of the image. This perspective model helps to create a very impressive illusion and allows a view of the theatre scene different from the one of Baroque era. The new system created a background with the vision by corner and not from the front of the building: in this way it was possible to have a better visibility from anywhere and a greater imagination of depicted space.

The paper, accompanied by graphic and photographic apparatus, reconstructs the spatial layout depicted in the painting of the curtain using photogrammetry, the practice of obtaining accurate measurements from a photograph to reconstruct the orientation of the geometric construction. This operation will be preliminary to a critical interpretation of the represented space to study the perception from different points of the theatre hall, but also to compare the allusive space proposed in the theatre setting (with particular reference to the staircase) and the real space of the main vestibule of the royal palace, as represented by Luigi Vanvitelli in the "*Dichiarazione dei Disegni*" of 1756 and subsequently realized. In fact there are several elements that link the image on the curtain to the built space, not only in the plant floor plan but also, for example, in the presence of the sculpture of Hercules as depicted on the canvas but also present in the central hall of the Royal Palace, at the basis of the staircase.

Keywords: perspective, scenography, illusion, Caserta.

1. INTRODUCTION

The aim of this paper is to trace, within the very most well-known Royal Palace of Caserta, a study of an element, present in the theater of the palace itself.

The Royal Palace was built by the King of Naples Carlo of Borbone who, struck by the beauty of the landscape of Caserta and eager to give a worthy representative office in the government of the capital city of Naples, wanted to be built a royal palace that it can hold its own

against the French Versailles.

As well known, for the project the King was entrusted to the architect-engineer Luigi Vanvitelli to whom he committed not only the Palace but also the park and the design of the surrounding urban area, with the procurement of a new aqueduct that arrive to the complex of San Leucio. The new palace was to be a symbol of the new state of Borbone House and prove the power and the grandeur of the family itself, but also be efficient and rational.

Among the elements that characterize the

whole palace is interesting to dwell on the theater, which opened in 1769, with an horse-shoe-shaped plant with five tiers of boxes, adorned with bundles twelve columns from the Temple of Serapis at Pozzuoli.



Figure 1: Luigi Vanvitelli, Tito Manlio scene



Figure 2: Luigi Vanvitelli, Monumental hall scenography

Recognized by scholars as one of the most significant examples of the theater, its constructive story is strongly linked to the one of the building that encompasses it.

Throughout the building Vanvitelli testifies his constructive attention that is also found in the theater for which the architect demonstrates not only his assiduous attendance of entertainment halls but also its participation in theatrical productions also found in numerous drawings of theatrical scenes [1]. (Figg. 1-2)

In the Theatre of Caserta, the architect creates an artifice to integrate the theater scene with the outside of the theater itself: the stage, almost as big as the stalls area, culminate in a great portal that opens to the side garden of the royal palace, giving rise to a perspective of great dramatic effect.

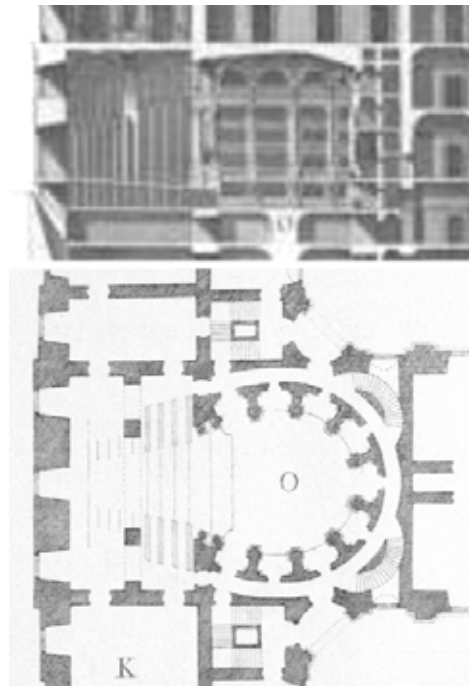


Figure 3: Luigi Vanvitelli, *Dichiarazione dei disegni del Real Palazzo di Caserta*. Royal theater: cross section and plan.



Figure 4: Royal Palace in Caserta, the curtain of the theatre (Photo by Luca Canonici).



Figure 5: Royal Palace in Caserta, theater. View from the Royal box to the stage.



Figure 6: Royal Palace in Caserta, theater. View from the stage to the stalls area.

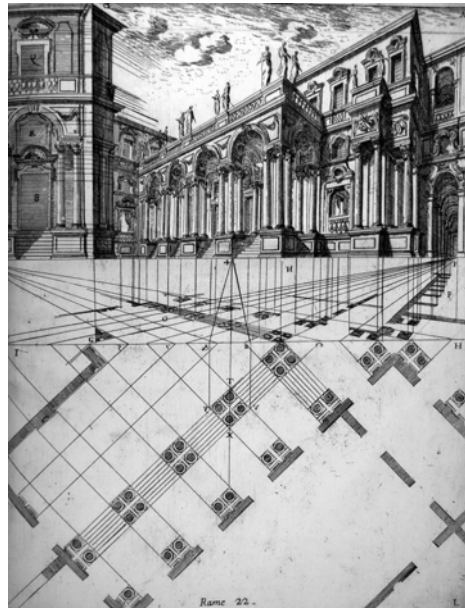
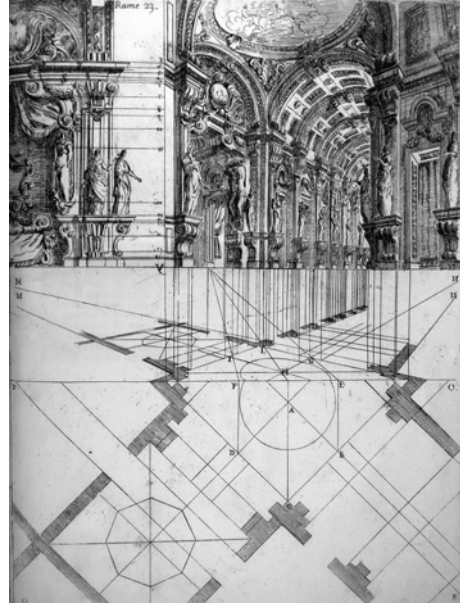
A careful that testifies to profound attention of Vanvitelli in the study of the built spaces of architecture and their integration with nature. Vanvitelli personally oversaw both the technical aspects relating to the construction of the theater, with the difficulties related to the construction of the great vault and with the proscenium arch, and all decorations of the great hall. Antonio Joli, however, was only entrusted with the planning and the execution of the six stage mutations as well as the construction of the machinery of the scene [2]. The scenic systems realized are characterized by some differences with those shown in the drawings of the *"Dichiarazione dei disegni del Real Palazzo di Caserta alle sacre reali maestà di Carlo Re delle Due Sicilie"* [3] (Fig.3) and with the drawing in the *Platea* by Sancio which does not mention any details. In fact, an inventory drawn up in 1846 by Fortunato Quériau registers the presence of 16 canvases on each side of the stage as well as a variety of additional equipment to support the needs of theatrical performances and illumination facilities for the scene and the creation of effects with the aid of the same light. Another element on the theater stage to be mentioned is the large canvas identifiable as curtain, not yet assigned by the recent historiography, attributed by some to Girolamo Starace Franchis already active in Caserta at Borbone court [4].

The painting, subject of this research, depicts an urban scene on two levels. Above a square enclosed by a semicircular colonnade concluded by a classical temple with pronaos, bottom a double flight staircase with balustrade, of which it is most visible only one. The scene is also enhanced by the characters and sculptures [5]. (Figg. 4-5-6)

2. METHODOLOGY

The perspective setting of the image depicted in the curtain fits fully to those who were the precepts of the time. Let us recall that the history of the representation theories and methods recognizes the determinate impulse given by Bibiena family to the innovative use of perspective in the eighteenth century. Martin Kemp says that the

last phase of the Italian perspective painting can not be considered apart from the innovative contribution of the Bibiena family of which at least eight members chose the way of scenic design [6]. In particular, Ferdinando was formed at the Accademia Clementina, which produced so many scenic designers and illusionists and wrote a treatise titled *Paradossi per praticare la prospettiva senza saperla*. In this work Ferdinando Galli Bibiena introduces significant innovations not only in terms of geometric constructions but also for the design of architectural elements. The major innovation introduced concerns the so-called '*scena per angolo*' (scene for corner) that is the real innovation introduced by Bibiena than the models used until then. In this design solution, later used also by Filippo Juvarra and other scenic designers, the central vanishing point (intersection between the line of sight with the picture plane) of perspective is placed on the edge or even outside the framework of the perspective scene, then out of the spectator's eyes, in a different location than the central position typical of the classical tradition. This artifice ensures a more dynamic space and therefore represented a strong illusionist suggestion. The illusory effect is further expanded by a bold use of proportional scales that draw the viewer into the scene represented. The scheme of '*scena per angolo*' sets the plan of the represented space in angular position, therefore in a completely different position from the one of the baroque scenes which favored the point of view of the King position to the detriment of other positions to which the scenic image was partially perceived [7-8]. The new setting of the scene instead ensured a better perception of the represented image by different points of view in the theater as well as a lower perception of the distortions and perspective aberrations from the viewpoints side. The final effect ensures a great illusory perception of the scene represented. A setting of the scene that exceeds the classical scheme of symmetry in favor of a scheme that would allow to a spectator in marginal position to perceive the same image of a real space [9-10]. (Figg. 7-8)



Figures 7-8: Ferdinando Galli Bibiena. Scenic design.

3. RESULTS

The knowledge on the represented space, aluded in the scene of the large canvas, is analyzed through photogrammetry. It is an application of descriptive geometry that through graphical operations exclusively give back to the 'true shape and dimension' of the objects represented by procedures opposite to the ones used for the construction of perspective, recognizing the biunivocal relation between designed space and real space [11].

This method also applied to photographic images, on the basis of the analogy between the latter and the projective perspective, was invented by the French Colonel Aimé Laussedat (1819-1907) who adopted it initially for topographical survey and later also to architecture. He was also the author of the treatise on the subject entitled "*Histoire de l'application de la photographie au levée des plans*".

In applying the above method to the image of the canvas is perceived the author's choice of identifying two perspective reference systems, in particular, one for the space close by the an hemicycle with columns and another one for the staircase located down to the previous mentioned space.

For both perspective reference systems, the central vanishing point is on the right edge of the image almost on the same vertical. The position of the view point or better the distance of these from the canvas changes in the two perspective representation. (Fig. 9)

Assumed the image as a perspective with vertical position of the picture plane, the first data obtained is the positions of the two central vanishing points V_{01} and V_{02} and the distances d_1 and d_2 that, with the horizon line (eye level), are the so-called 'inner orientation'. From these it is possible to redraw the plant and the elevation of the depicted space both of the higher one -place- (Fig. 10) than the lower one -staircase- (Fig. 11) and from these redraw in orthographic projections all the elements through the assembly of these two parties. (Fig. 12)

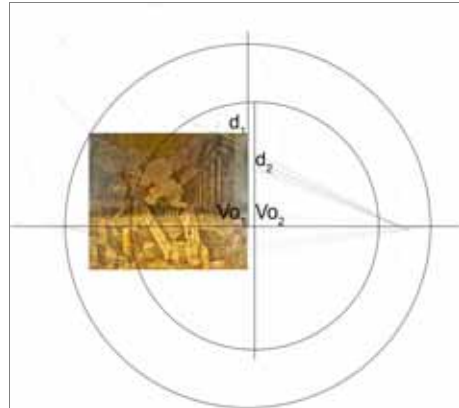
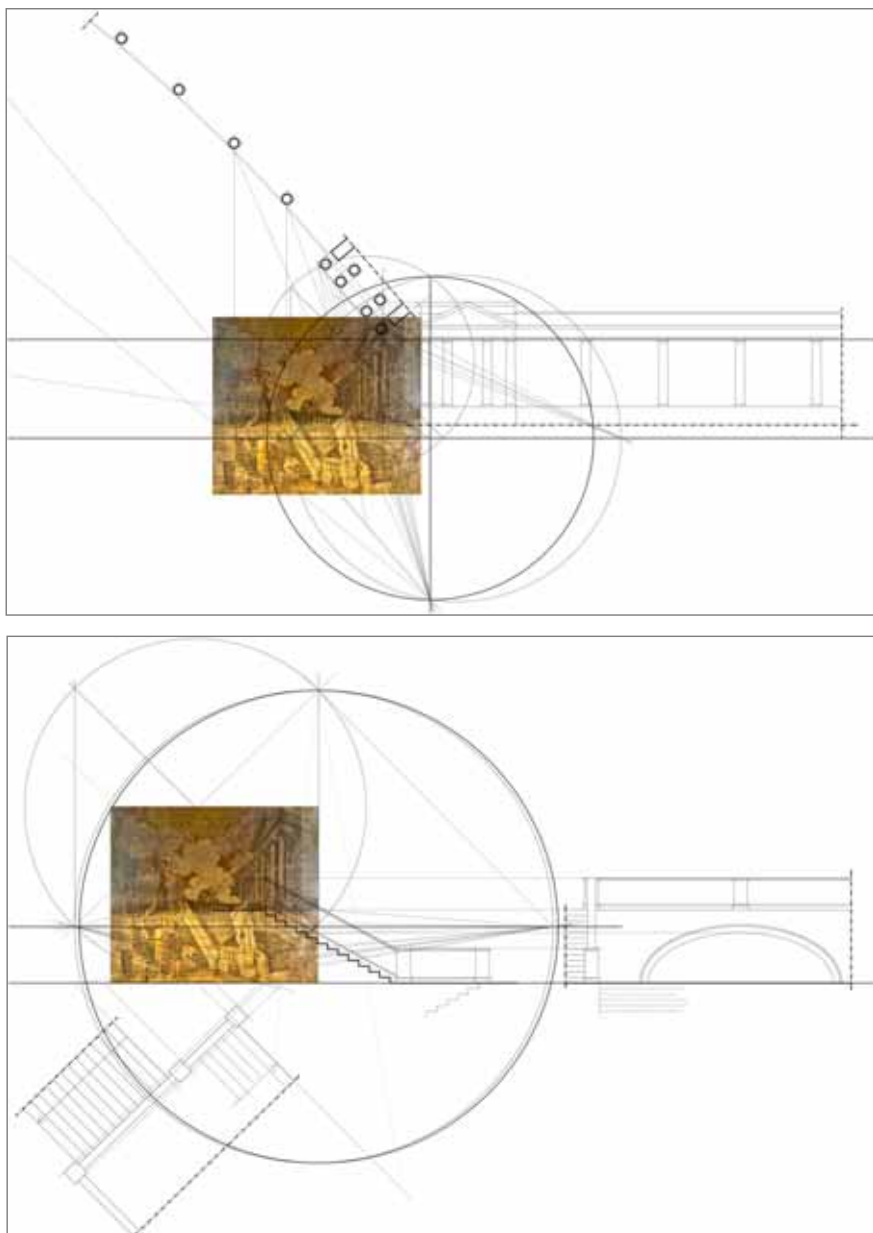


Figure 9: Royal Theater in Caserta: reference systems in the curtain image representation.

This operation was necessary in order to verify the assumptions that this study would verify: the perception of the canvas inside the Royal Theater and the analogy between the space represented in the canvas and the staircase of the same Royal Palace of Caserta. [12]

The verification of the perception of the image of the canvas from the inside of the theater was carried out by checking the positions of the centers of perspective view inside the theater. In particular, through a comparison between the current position of the canvas and the one that originally, as evidenced by the sources, might have as curtain of the stage. Well, the comparison points out that the current position (justifiable for logistical requirements due to a contemporary use of the stage for different functions than the theatrical one; the actual position is justifiable to leave the stage free for activity different from the theatrical performances) puts the two points of view V'_{1a} V'_{1b} , respect to which the image has been drawn, on the stage itself (Fig.13), while the position of the canvas as a stage curtain puts the viewpoints V'_{2a} and V'_{2b} inside the stalls area, in a better position for spectators. (Fig.14)



Figures 10-11: Royal Theater in Caserta: photogrammetry of the higher and lower parts of the curtain.

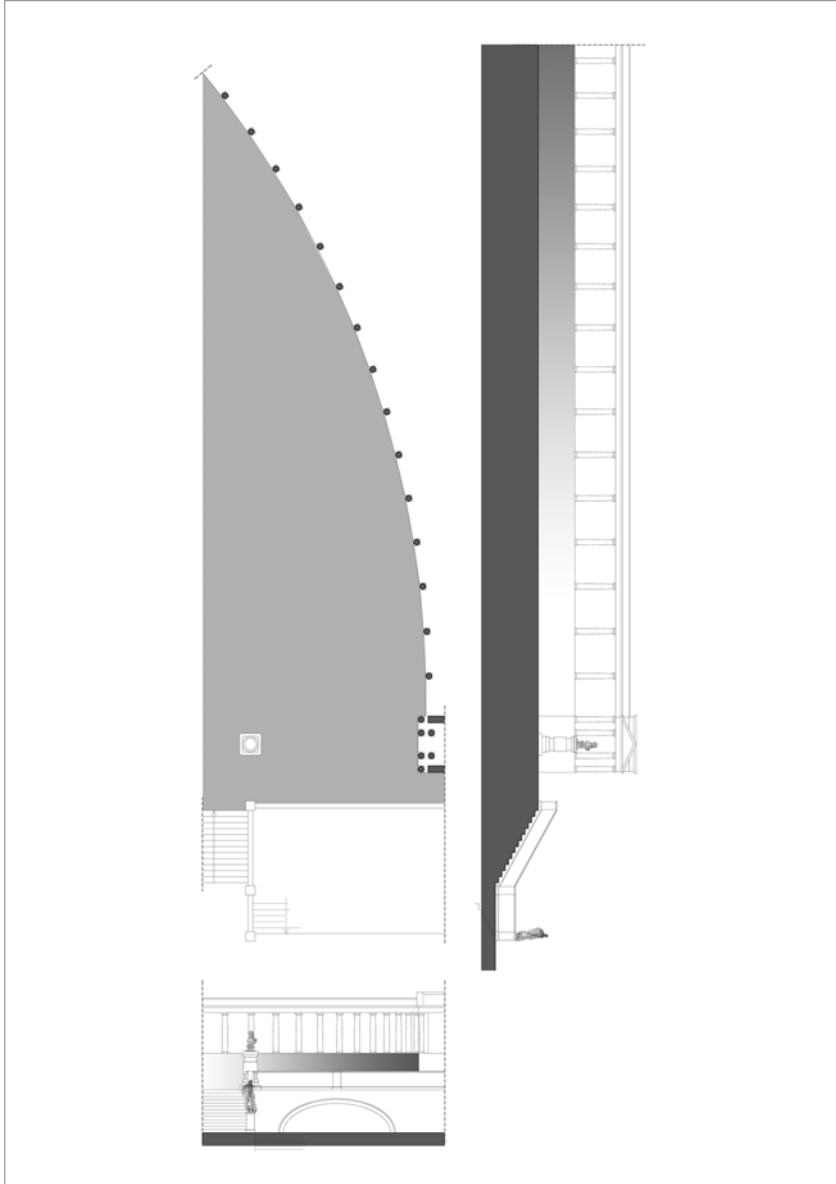


Fig. 12: Orthographic projection of the imagined space.

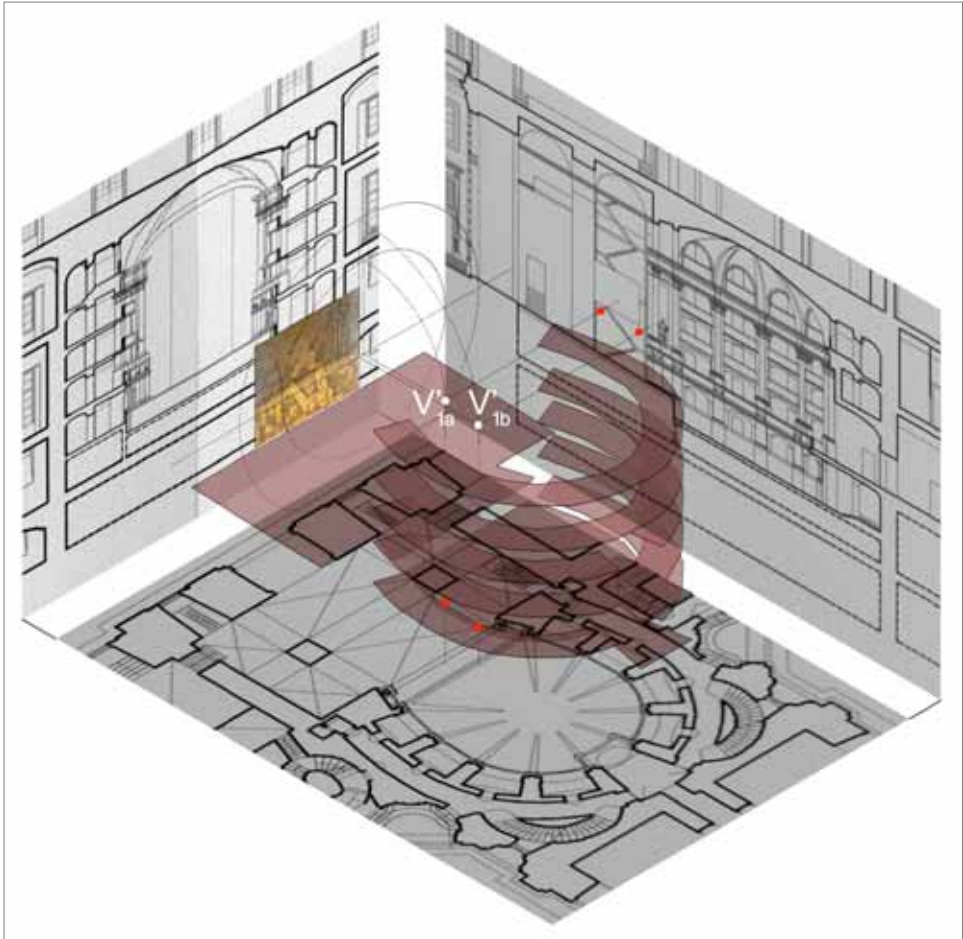


Fig. 13: Royal Theater in Caserta: the actual position of the canvas and the two points of view.

It's also true, as mentioned previously in regard to the '*scena per angolo*' setting, that an angular perspective, whether painted or solid perspective, allows a better perception of the scene itself to almost all positions of the theater. It is important to specify, however, that the horizontal plane through the point of view of the image of the great canvas of the Theatre of the Royal Palace lies at the altitude of a spectator seated at the level of the Royal box, thus en-

suring an optimal vision at that altitude. The setting of the perspective structure, realized according to the directions followed by Bibiena in their theatrical scenes, thus ensuring a dynamic of illusory image, through an angular position of the viewpoint, offers to spectators at the same altitude of Royal box and then to the real Sovereign himself the chance to enjoy the image of the curtain from an optimal point of view [13-14].

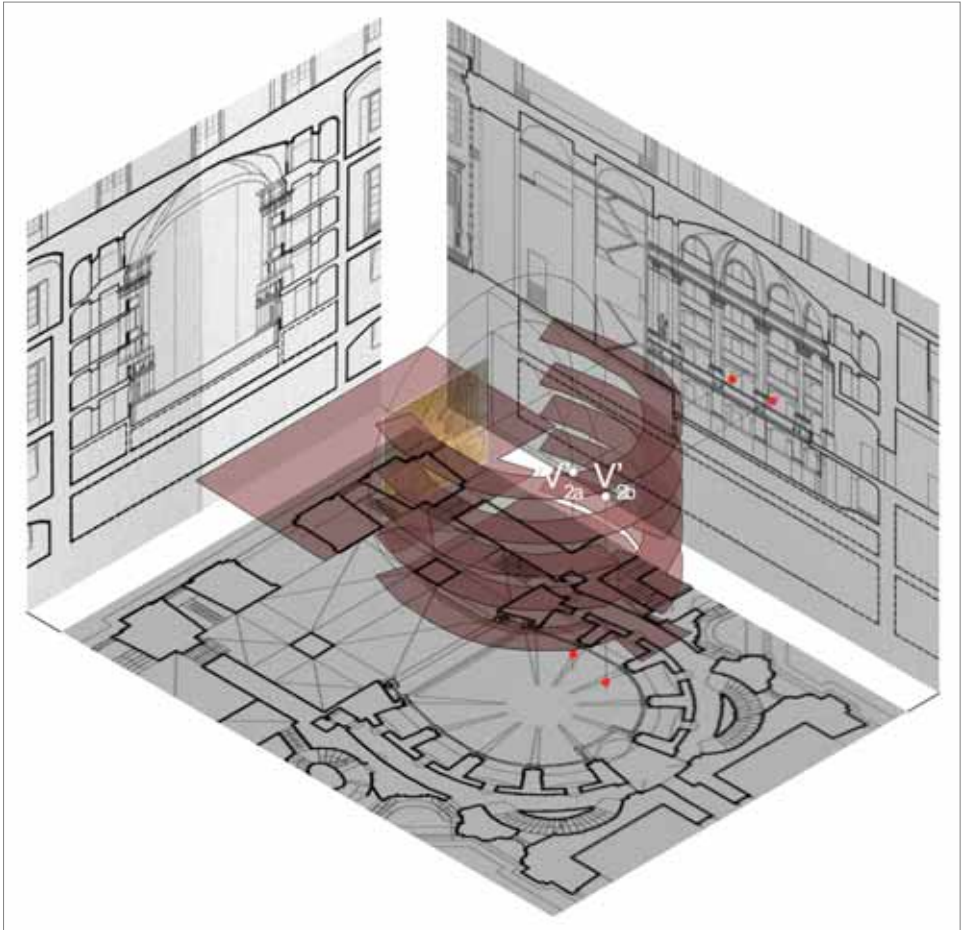


Fig. 14: Royal Theater in Caserta: the original position of the canvas and the two points of view.

A final observation should be developed in reference to what some authors claim about the analogy between the illusory space represented in the curtain of the Theatre of the Royal Palace in Caserta and the built space in the same building. Ciapparelli, in fact, in the description of the depicted image and, in particular, in connection with the presence in the foreground of the figure of Hercules placed on the balustrade of the staircase, but also present in the ground floor central hall of the Royal Palace,

asserts: "The reproduction of the famous Farnese sculpture is perfectly consistent with the symbolism of Borbone House. There were highlighted the connections between the myth of Hercules and the figure of Carlo III, due to the identification of the King with the mythical hero [...] Original is instead the idea of playing in the canvas a scenic space allusive to the central hall of the Royal Palace, where the Hercules by Andrea Violani introduces the staircase facing the statue of the "*Maestà Regia*" (Royal

Majesty) by Tommaso Solari on the first landing. Similarly, the Hercules painted on canvas introduces the flight of the staircase surmounted by a statue in whose it is possible to recognize the effigy of Carlo III" [4]. The formal analogy between the image, specifically the bottom part of the curtain, and the staircase is clearly supported from photogrammetry previously discussed: the arch surmounted by a balustrade is above the first flight of the staircase in both cases, the difference is the number of flights that in the perspective are two, while in reality the ramp that starts from the inner vestibule, came to the first landing, is divided into two parallel flights that reach both the upper vestibule. Instead, it is an evident analogy between the positions of the sculptures of Hercules and the King which in both cases are aligned along the same directrix of the scale with the King in a position of higher altitude than the one of the hero: in the imagined space Hercules is placed on the balustrade that separates the two ramps and a bust of Carlo III on the pedestal at the intersection of the axis of the balustrade and the axis of the pronaos of the semicircular colonnade, in real space Hercules and the Royal Majesty confront, looking each other, between the vestibule along the central gallery on the ground floor of the Palace and the first landing of the Royal staircase. (Figures 15-16-17)

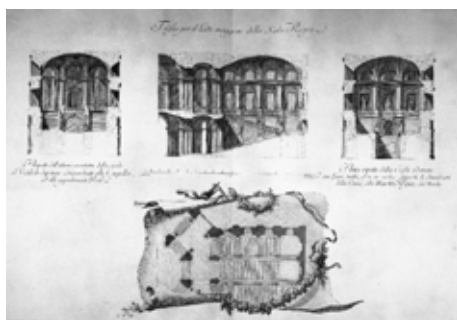


Fig. 15: Luigi Vanvitelli, *Dichiarazione dei disegni del Real Palazzo di Caserta*. Royal staircase: cross sections and plan.



Fig. 16: Royal Palace in Caserta: staircase. The sculpture of Hercules in the vestibule along the central gallery on the ground floor of the Palace



Fig. 17: Royal Palace in Caserta: staircase. The statue of "Maestà Regia" on the first landing.

4. CONCLUSIONS

This paper aims to offer a contribution to knowledge, based on the use of the methods of descriptive geometry, to the study of a lesser known element of the Royal Palace in Caserta, protected by UNESCO World Heritage List, in order to contribute to the research and enhancement also through the study of integration between science and art, in the period in which the palace was built, finalized to the material and cultural safeguard of the canvas.

ACKNOWLEDGMENTS

The author would like to thank Luca Canonici for the orthometric picture of the canvas in the Theatre of the Royal Palace in Caserta and Antonella Diana of "Soprintendenza Speciale per il Patrimonio Storico Artistico ed Etnoantropologico e per il Polo Museale della Città di Napoli e della Reggia di Caserta" Office.

REFERENCES

- [1] Claudio Marinelli (a cura di), *L'esercizio del disegno. I Vanvitelli. Catalogo generale del fondo disegni della Reggia di Caserta*, Leonardo De Luca Editore, Roma 1991
- [2] Soprintendenza per i beni ambientali, architettonici, artistici, e storici per le province di Caserta e Benevento, *Il Teatro di Corte di Caserta: storia e restauro*, Electa, Napoli 1995.
- [3] Luigi Vanvitelli, *Dichiarazione dei disegni del Real Palazzo di Caserta alle sacre reali maestà di Carlo Re delle Due Sicilie...*, Napoli 1756
- [4] Pier Luigi Ciapparelli, *Luigi Vanvitelli e il teatro di corte di Caserta*, Electa, Napoli 1995.
- [5] Vega De Martini, Francesco Petrucci (a cura di), *Vanvitelli segreto. I suoi pittori tra Conca e Giaquinto, la "Cathedra Petri"*, Gangemi Editore, Roma 2014.
- [6] Martin Kemp, *La scienza dell'arte. Prospettiva e percezione visiva da Brunelleschi a Seurat*, Giunti Editore, Firenze 1994.
- [7] Andrea Giordano, *La geometria nell'immagine. Dal secolo dei Lumi all'epoca attuale*, UTET, Torino 2002.
- [8] Mary L. Myers, *Architectural and Ornament Drawings. Juvarra, Vanvitelli, the Bibiena Family, and Other Italian Draughtsmen*, Published by The Metropolitan Museum of Art/Distributed by Yale University Press, 2013.
- [9] Fauzia Farneti, Deanna Lenzi (a cura di), *L'architettura dell'inganno. Quadraturismo*

e grande decorazione nella pittura di età barocca, Alinea Editrice, Firenze 2004.

- [10] Giannino Paliaga (a cura di), *Il Gran Teatro delle Fabbriche. Libri di architettura dal XV secolo al XIX secolo*, Biblioteca di via Senato Edizioni, Roma 2000.
- [11] Mario Giovanardi, *Elementi di fotogrammetria applicata all'architettura*, Napoli 1958.
- [12] Cesare Cundari, *Il Palazzo Reale di Caserta*, Edizioni Kappa, Roma 2005.
- [13] Alessandra Pagliano, *Il disegno dello spazio scenico. Prospettiva illusoria ed effetti luminosi nella scenografia teatrale*, Hoepli, Milano 2002.
- [14] Giorgio Ricchielli, *La rappresentazione prospettica e il progetto scenografico*, Cluva Città Studi, Venezia 1991

ABOUT THE AUTHOR

Nicola Pisacane: Assistant Professor of Drawing - Department of Architecture and Industrial Design Luigi Vanvitelli - Seconda Università di Napoli. PhD in Environmental and Architecture Surveying and Imaging - Seconda Università di Napoli. Member of PhD School in Architecture, Industrial Design and Cultural Heritage - Seconda Università di Napoli. Member of BENECON SCaRL Research Center on Cultural Heritage Ecology Economy. Member of research project and author of scientific papers and proceedings of international and national congresses about landscape, territorial and architectural representation, geometry.
nicola.pisacane@unina2.it

IMPROVEMENT OF QUADRIC ERROR METRIC MESH SIMPLIFICATION BY REEB-GRAPH-BASED TOPOLOGICAL INFORMATION

Chinnawat DEVAHASTIN NA AYUDHYA and Pizzanu KANONGCHAIYOS
Chulalongkorn University, Thailand

ABSTRACT: This paper presents an enhancement to quadric error metrics (QEM) in order to preserve appendages of models with high topological complexity while they are simplified by adding weights obtained from the Reeb graphs of the models. For precomputation, the integral of geodesic distance is calculated at each vertex in the same way as constructing the Reeb graph. The calculated values are divided by the maximum value to get the normalized values, $\mu(v)$. To compute initial QEM matrices, each face is weighted with $\mu(f)10$ where $\mu(f)$ is the average of $\mu(v)$ of face vertices. When scoring each edge, the raw score from QEM is divided by $\mu(e)6$ where $\mu(e)$ is the average of $\mu(v)$ of the edge's endpoints. The four-degree difference allows more faces to preserve appendages. Experiments are conducted with six models on Intel Core i7-4700HQ 2.40 GHz with 8.00 GB RAM. The results show that the proposed algorithm can keep significant topological features of the models. In addition, the Hausdorff distance of the proposed algorithm is lower than that of the original QEM when models start to degenerate. For meshes of which the numbers of vertices range from 48,485 to 314,218, the decimation time of the proposed algorithm is approximately 147.91 microseconds per vertex, which is almost equal to the decimation time of QEM which is 145.32 microseconds per vertex. The additional time to compute the integral of geodesic distance for the same set of meshes is roughly 146.26 microseconds per vertex on average, which is reasonable for the better results.

Keywords: Mesh simplification, edge contraction, quadric error metric, Reeb graph, topological information.

1. INTRODUCTION

3D models have been increasingly used in games, applications, and animations to simulate a satisfying, real-looking environment. A 3D model is usually represented by a mesh which consists of connected triangles covering the model's surface. To obtain a structure, several approaches have been utilized, such as using a 3D scanner and digital sculpting. However, most of the methods generate a high-resolution model, which reduces the model usability. A large model consumes processing time, drawing time, and memory. Especially for applications on mobile devices which have bounded CPU speed, limited memory, and constrained

energy, it is infeasible for a gigantic mesh to be exploited on such hardware. Mesh simplification is introduced to cope with the problems. It reduces vertices and faces of the models while keeping their looks. Mesh simplification can also be used implicitly; for example, to compare 3D models, instead of comparing two big models, simplified models are used as representatives to save analyzing time.

2. RELATED WORK

The widely-used mesh-simplifying method is edge contraction where an edge with the least effect to the model is collapsed into a single vertex. After scoring each edge, the method utilizes a heap to find the least-scored edge,

and the model and the heap are gradually updated during the decimation.

One of the fast and effective algorithms to select edge candidates is quadric error metrics [2] proposed by Garland and Heckbert. A QEM matrix is associated with each vertex to carry additive squared distances from original faces, which can be viewed as planes, in order that, after an edge contraction, the position of the target vertex will minimize the squared distances between the target vertex and neighboring planes before the decimation. However, the algorithm tends to give rough results when used to decimate models with high topological complexity including limbs, appendages, and sharp corners. Because QEM matrices of an edge's endpoints are added to acquire the QEM matrix of the target vertex, at a sharp corner, the target vertex will move away from the corner to minimize the cumulative distance from the larger part.

In later work, principal curvatures [4] are included in QEM scores with expectation that moving a vertex on a flat surface should effect the model's structure less than moving a vertex in a high-curvature direction. Nevertheless, it is stated in the paper that the curvature-related factors are impractical when the model is reduced into a small size.

For topological representation, several studies [1] [3] have effectively used the Reeb graph to represent the model's topology because of its robustness to affine transformations. The integral of geodesic distance from each vertex to any other vertex is computed to obtain a value to segment the model into parts represented by Reeb nodes. Full computation of the integral of geodesic distance exploits high computational resources; therefore, base vertices [3] are randomly selected, and only shortest paths to base vertices are calculated.

3. METHODOLOGY

The proposed algorithm starts with computing integral geodesic distance to obtain $\mu(v)$ for each vertex v . $\mu(v)$ is then used to initialize

QEM matrices and score each edge. The edge with the least score is selected to be collapsed using a heap, and the model and the heap are updated after an operation. The process iterates until the preferred size is reached. The algorithm in details is described in this section.

3.1 Integral geodesic distance computation

The scheme to find the integral of geodesic distance is the same as what was used in creating the Reeb graph. The algorithm [3] utilizes Dijkstra's algorithm as summed up in Figure 1, where $r = \sqrt{0.005 \cdot \text{area}(\text{Surface})}$. $\mu(v)$ for each vertex v is computed by dividing $ig(v)$ by the maximum integral geodesic distance.

Algorithm 1	Integral geodesic distance computation
Input	A mesh
Output	The same mesh with an integral geodesic distance value $ig(v)$ at each vertex
Begin	
01:	For each vertex v , set it as unvisited, $g(v) = \infty$, $ig(v) = 0$, and $\text{area}(v) = 0$.
02:	While there is an unvisited vertex
03:	Randomly choose an unvisited vertex to be a base vertex v_b and put it into a heap.
04:	While the heap is not empty
05:	Pop a vertex v from the heap, mark it as visited, and set $\text{base}(v) = v_b$.
06:	For each vertex v_a adjacent to v , if $g(v) + \text{length}(v, v_a) < g(v_a)$, set $g(v_a) = g(v) + \text{length}(v, v_a)$, and if $g(v_a) < r$, put v_a into a heap.
07:	End While
08:	End While
09:	For each face $f(v_1, v_2, v_3)$, set $\text{area}(\text{base}(v_1)) = \text{area}(\text{base}(v_1)) + \text{area}(f)$.
10:	For each base vertex v_b , use Dijkstra's algorithm to obtain shortest paths to any other vertex v , set $ig(v) = ig(v) + \text{ShortestPath}(v_b, v) \cdot \text{area}(v_b)$.
End	

Figure 1: The integral geodesic distance computation algorithm

3.2 QEM matrix initialization

To obtain an initial QEM matrix $Q(v)$ for each vertex, the fundamental error quadric for each face, K_f , is weighted with $\mu(f)^{10}$ as shown in Equation (1), where $\mu(f)$ is the average of $\mu(v)$ of face vertices. As same as the original approach, the matrix K_f is constructed as depicted in Equation (2). Constants a , b , c , and d represent the face's plane described by $ax + by + cz + d = 0$ where $a^2 + b^2 + c^2 = 1$.

$$Q(v) = \sum_{f \in \text{AdjacentFace}(v)} K_f \cdot \mu(f)^{10} \quad (1)$$

$$K_f = \begin{bmatrix} a^2 & ab & ac & ad \\ ab & b^2 & bc & bd \\ ac & bc & c^2 & cd \\ ad & bd & cd & d^2 \end{bmatrix} \quad (2)$$

In order to preserve the model's boundaries, for each border edge, a simulated plane is created through the edge perpendicular to the face the edge belongs to. The fundamental error quadric for the special plane is weighted with $\mu(f)^{10}$ where f is the edge-hosting face.

3.3 Edge scoring

To acquire a score for collapsing an edge, the QEM matrices of the edge's endpoints are added to be the QEM matrix of the target vertex. Similar to the original approach, the position of the target vertex is calculated using matrix inversion. If the matrix is not invertible, the average of the positions of the edge's endpoints is used as the target position. The raw QEM score for each edge is divided by $\mu(e)^6$ where $\mu(e)$ is the average of $\mu(v)$ of the edge's endpoints. This helps prevent very dense faces at appendages.

3.4 Heap update

A min heap is utilized to select an edge with the least score, and each heap node represents

an operation to collapse an edge. A node keeps EdgeID, EdgeScore, and EdgeVersion. EdgeID is used to link the operation to an edge on the model. EdgeScore is used to store the score which is the heap key. EdgeVersion is used to check whether the operation is obsolete. If the operation is obsolete, the node is just popped out of the heap without simplification action. For each edge, CurrentEdgeVersion is stored. If EdgeVersion of the top operation is not equal to CurrentEdgeVersion of the edge, the node is considered out-of-date.

Algorithm 2 Overall mesh simplification	
Input	A mesh
Output	A lower-resolution mesh
Begin	
01:	Calculate the integral of geodesic distance $ig(v)$ for each vertex and normalize the values.
02:	Initialize a QEM matrix $Q(v)$ at each vertex.
03:	Call UpdateEdge function for every edge on the mesh.
04:	While the preferred size is not reached
05:	Pop an edge operation with the least score out of the heap where EdgeVersion equals CurrentEdgeVersion.
06:	Collapse the edge.
07:	Interpolate $\mu(v)$ for the target vertex.
08:	Call UpdateEdge function to update all edges adjacent to the neighboring vertices of the edge's endpoints.
09:	End While
End	

Figure 2: The overall mesh simplification algorithm

UpdateEdge function is created to update the new score of an edge. First, it increases CurrentEdgeVersion and checks whether the edge can be collapsed including that the normal vectors of faces adjacent to exactly one endpoint cannot change more than 90 degrees after the operation. Otherwise, face flips can occur. If the operation passes all criteria, a new heap node is created. EdgeScore is set to be the new score of the edge, and EdgeVersion is set to be

CurrentEdgeVersion. The node is then pushed into the heap.

To initialize the heap, UpdateEdge is called for every edge on the model. After an edge is collapsed, $\mu(v)$ of the target vertex is linearly interpolated. UpdateEdge is called to update all edges adjacent to the neighboring vertices of the edge's endpoints. The overall algorithm is shown in Figure 2.

4. EXPERIMENTS AND RESULTS

The proposed algorithm was implemented with OpenMesh 2.4 and tested on six models with high topological complexity including Horse, Armadillo, Tiger Fighter, Orc Hand, Dragon Head, and Dinosaur. The integral of geodesic distance of each model is illustrated in Figure 3 where green indicates low value and pink indicates high value.

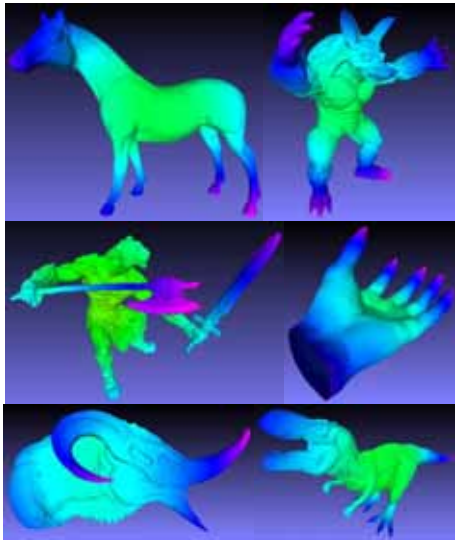


Figure 3: The integral of geodesic distance

An enhanced QEM which includes prevention of boundary corrosion and face flips was implemented with OpenMesh 2.4 to be compared with the proposed algorithm. The results show that the proposed algorithm can preserve

topological features of the models better than QEM. Some of the results are depicted in Figure 4 to Figure 7. Although Orc Hand has a boundary, the proposed algorithm can preserve both its boundary and its features.

Moreover, the Hausdorff distance of the proposed algorithm is lower than that of the QEM when models start to degenerate. Provided that A and B are sets of vertices of an original mesh and a simplified mesh respectively, the Hausdorff distance, $H(A, B)$, is the maximum of $h(A, B)$ and $h(B, A)$ where $h(A, B)$ is the maximum among distances from each vertex in A to its nearest vertex in B. $h(A, B)$ and $h(B, A)$ were computed by MeshLabServer. The Hausdorff distances relative to the diagonal of the model's bounding box of all six models are shown in Figure 8 to Figure 13.

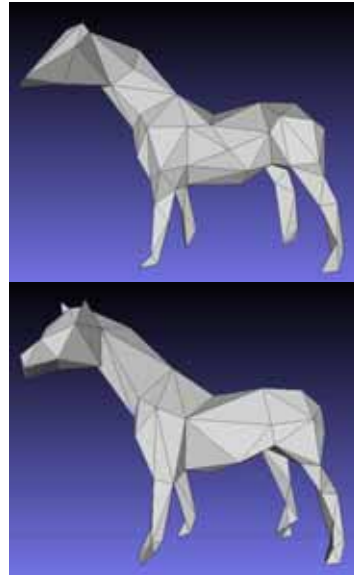


Figure 4: 0.2% Horse simplified by QEM (top) and by the proposed algorithm (bottom)

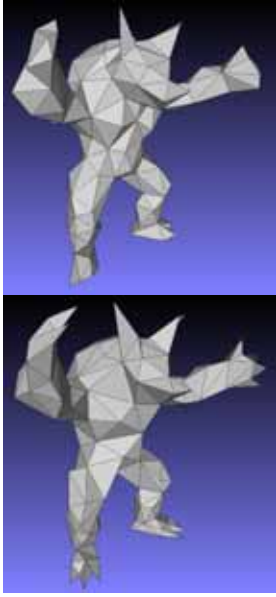


Figure 5: 0.1% Armadillo simplified by QEM (top) and by the proposed algorithm (bottom)

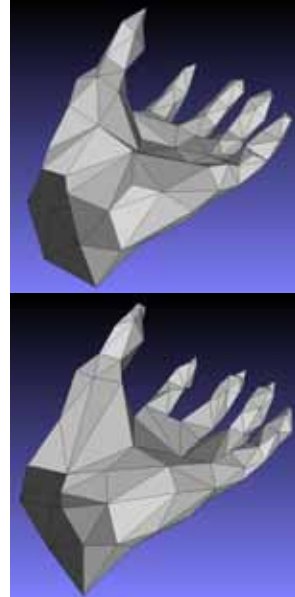


Figure 7: 0.05% Orc Hand simplified by QEM (top) and by the proposed algorithm (bottom)

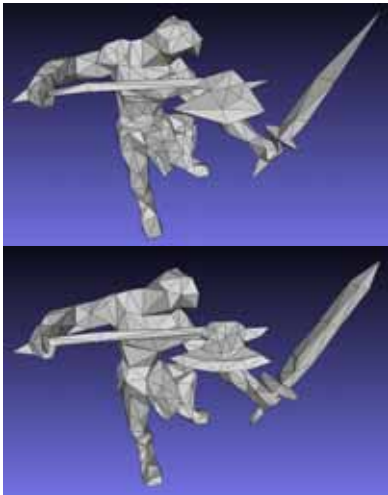


Figure 6: 0.1% Tiger Fighter simplified by QEM (top) and by the proposed algorithm (bottom)

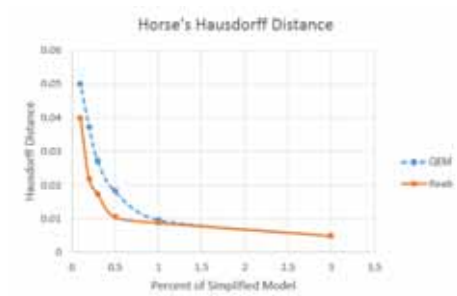


Figure 8: Horse's Hausdorff distance

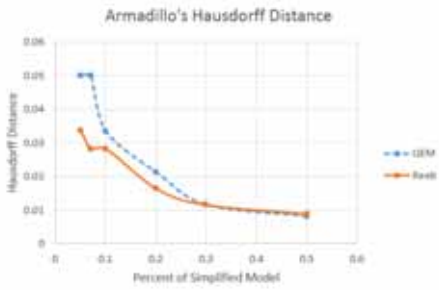


Figure 9: Armadillo's Hausdorff distance

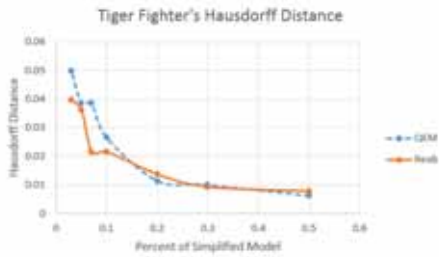


Figure 10: Tiger Fighter's Hausdorff distance



Figure 11: Orc Hand's Hausdorff distance

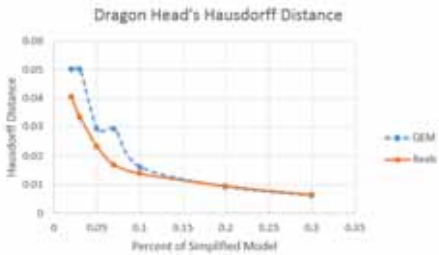


Figure 12: Dragon Head's Hausdorff distance

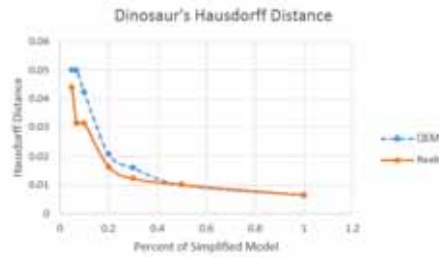


Figure 13: Dinosaur's Hausdorff distance

Table 1: Calculation time

Model Name	No. of Vertices	No. of Base Vertices	IGD Comp Time (s)	Decimation Time (s)
Horse	48,485	185	5.539	6.668
Armadillo	172,974	194	25.422	26.534
Tiger Fighter	314,218	186	42.990	47.046
Orc Hand	275,617	243	43.749	38.076
Dragon Head	211,404	214	38.556	33.149
Dinosaur	149,984	199	19.865	22.213

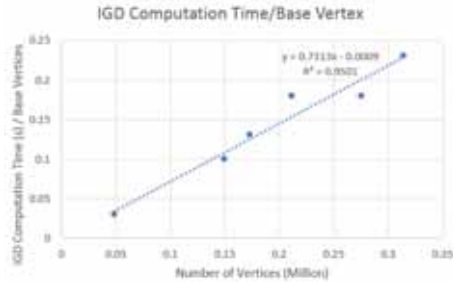


Figure 14: The integral of geodesic distance computation time per base vertex

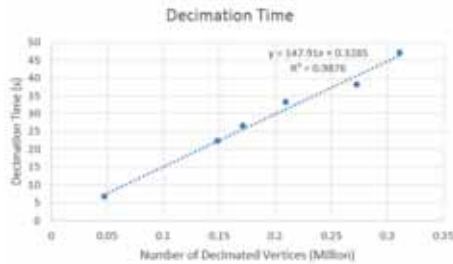


Figure 15: Decimation time

For calculation time, each model was simplified to 1% of its vertices on Intel Core i7-4700HQ 2.40 GHz with 8.00 GB RAM. The time to calculate the integral of geodesic distance (Algorithm 2 Step 1) and the time to perform a simplification (Algorithm 2 Step 2-9) are shown in Table 1. The amount of time to calculate the integral of geodesic distance per base vertex is approximately 0.7313 microsecond per vertex as shown in Figure 14. If a model has 200 base vertices, the time will be 146.26 microseconds per vertex. The simplification time is quite linear at roughly 147.91 microseconds per vertex as shown in Figure 15, which is almost equal to the decimation time of QEM which is 145.32 microseconds per vertex.

5. DISCUSSIONS

The proposed algorithm can preserve the topological features of a model when it starts to degenerate; however, the results of the algorithm at high resolutions are quite the same as the results of the original QEM. Therefore, a model can be first decimated by QEM and then simplified by the proposed algorithm at later stages. It should be noted that if the model is too rough before computing the integral of geodesic distance, errors of $\mu(v)$ will dramatically rise because the vertex distribution cannot represent the continuity of $\mu(v)$.

Although Hausdorff distance is a great metric to judge similarity between models, it cannot reflect much about topology preservation of a model. A model with fully-preserved appendages can bring about higher Hausdorff distance than a model with eroded appendages. Hence, another metric focusing on topology preservation should be further studied.

Not all kinds of models fit the proposed algorithm. A model that perfectly suits the algorithm should have appendages, and the appendages should be key features of the model. More importantly, the integral of geodesic distance must reflect the appendages of the model; otherwise, the weights are meaningless. The

following example of Wooden Chair, shown in Figure 16, can depict advantages and disadvantages of the algorithm.



Figure 16: Wooden Chair with the integral of geodesic distance



Figure 17: 0.1% Wooden Chair simplified by QEM (left) and by the proposed algorithm (right)

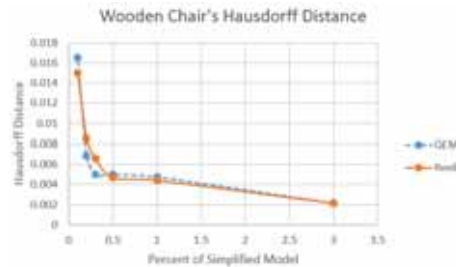


Figure 18: Wooden Chair's Hausdorff distance

In Figure 17, it is obvious that the proposed algorithm works better on the bottom-half of

the model while QEM works better on the top-half of the model. It is because, in the bottom-half, legs which are appendages of the model are reflected properly by the integral of geodesic distance. On the other hand, in the top-half, there are no appendages, and the integral of geodesic distance cannot reflect much about the model's features. Thus, the proposed algorithm merges the seat with the back slats. Figure 18 shows that the proposed algorithm cannot improve the Hausdorff distance of the model. A solution is stated in the next section.

6. CONCLUSIONS

This paper describes an improvement to quadric error metric mesh simplification in order to preserve topological features of models. Information from the Reeb graph of the model, which is the integral of geodesic distance, is used to weight initial QEM matrices and edge scores. Boundary corrosion and face flips are prevented in both QEM implementation and the proposed algorithm implementation. The results indicate that the proposed algorithm can keep appendages and result in lower Hausdorff distance than QEM at later stages. The decimation time of the proposed algorithm is almost equal to that of QEM, and the additional time to compute the integral of geodesic distance for meshes of which the number of vertices are between 48,485 and 314,218 is approximately 146.26 microseconds per vertex.

For future work, local weights which describe local features of a model should be added together with the global weights, which can be acquired from the integral of geodesic distance to preserve both local and global features of the model. In addition, for a better usage of the proposed algorithm, a model should be segmented so that each part has appendages as features. For example, for the Wooden Chair, if the model is divided into two parts, the top-half and the bottom-half, each part will suit the proposed algorithm perfectly.

ACKNOWLEDGMENTS

This research is supported by Junior Science

Talent Project by the National Science and Technology Development Agency, Thailand. Tiger Fighter, Dinosaur, and Wooden Chair models are obtained from the gallery of scanned 3D models using Artec 3D scanners (<http://www.artec3d.com/gallery/3d-models/>).

REFERENCES

- [1] W. Areevijit and P. Kanongchaiyos. Reeb graph based partial shape retrieval for non-rigid 3D object. In *Proceedings of the 10th International Conference on Virtual Reality Continuum and Its Applications in Industry*, pages 573-576, 2011.
- [2] M. Garland and P. S. Heckbert. Surface simplification using quadric error metrics. In *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques*, pages 209-216, 1997.
- [3] M. Hilaga, Y. Shinagawa, T. Kohmura, and T. L. Kunii. Topology matching for fully automatic similarity estimation of 3D shapes. In *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques*, pages 203-212, 2001.
- [4] V. Ungvichian and P. Kanongchaiyos. Quadric-based mesh simplification method with penalty based on principal curvatures and directions. *International Journal of Computer Science Issues*, Vol. 8, Issue 4, 2011.

ABOUT THE AUTHORS

1. Chinnawat Devahastin Na Ayudhya received a high school diploma in science and mathematics from Mahidol Wittayanusorn School. He is currently studying for his B.Eng. degree at Chulalongkorn University.

2. Pizzanu Kanongchaiyos received a B.S. degree from Thammasat University, an M.S. degree from the University of Tokyo, and a Ph.D. degree from the University of Tokyo. He is currently an assistant professor in the Department of Computer Engineering, Chulalongkorn University.

THE INCIRCLE HYPERBOLOID AND ELLIPSOID OF THE CONJUGATE PAVILLET TETRAHEDRA

Axel PAVILLET

Singapore University of Technology and Design, Singapore

ABSTRACT: The Pavillet Tetrahedron is a unique orthocentric tetrahedron attached to a triangle called base triangle. Its has numerous properties which can be used to prove triangle geometry theorems or, conversely, triangle geometry theorems can be used to prove its properties. It is built from the incircle of the base triangle; the incenter is called the apex of the tetrahedron, the other three vertices form a triangle called the upper triangle. In this paper we consider the incircle as given. Hence the apex of the tetrahedron is fixed and the vertices of the upper triangle, which positions are a function of the tangential distance of the vertices of the base triangle to the incircle, describe a surface in \mathbb{R}^3 . While the vertices of the base triangle describe the plane of the incircle outside the incircle, the corresponding Gergonne point of this triangle describe the inner part of the incircle. Similarly, the orthocenter of the tetrahedron which project on the plane of the incircle as the Gergonne point of the base triangle describe a surface. We show that the vertices of the upper triangle lie on a one sheet equilateral hyperboloid of revolution while the orthocenter lies on an oblate ellipsoid of revolution inscribed in the throat circle of this hyperboloid. Moreover by symmetry two Pavillet tetrahedra can be built from a single base triangle. We call them conjugate tetrahedra because we show that the links between these two tetrahedra created by the hyperboloid is stronger than a simple symmetry.

Keywords: Tetrahedron, Hyperboloid, Ellipsoid, Revolution, Gergonne, Orthocenter

1 Introduction.

We recall that the orthocentric tetrahedron of a scalene triangle [4], named the Pavillet tetrahedron by Richard Guy [1, Ch. 5] and Gunther Weiss [7], is formed by drawing from the vertices A , B and C of a triangle ABC on the horizontal plane, called the base triangle three vertical segments $AA' = AM = x$, $BB' = BK = y$, $CC' = CL = z$ where KLM is the contact triangle of ABC . We then consider the tetrahedron $IA'B'C'$ (fig. 1). The three points A' , B' , C' form a triangle called the upper triangle and define a plane, the upper plane. The figure is rich in properties and various of them have been described in [4],[3] and [5].

We can build on both sides of the plane of the base triangle so the Pavillet tetrahedron of ABC , $IA'B'C'$ has a symmetric tetrahedron; we des-

ignate the second set of vertices and points by \bar{A}' , \bar{B}' , \bar{C}' and call this pair of tetrahedra the conjugate tetrahedra of the triangle. The orthocenters of the tetrahedra will respectively be called H'' and \bar{H}'' .

2 The Hyperboloid of revolution.

The nine points A' , B' , C' , $\bar{A}'\bar{B}'\bar{C}'$ and K , L , M define a unique quadric surface.

Theorem 2.1 *The quadric surface defined by the six vertices of the upper triangles of the conjugate tetrahedra and the contact triangle of the base triangle is a one sheet equilateral hyperboloid of revolution centered at I with a vertical axis.*

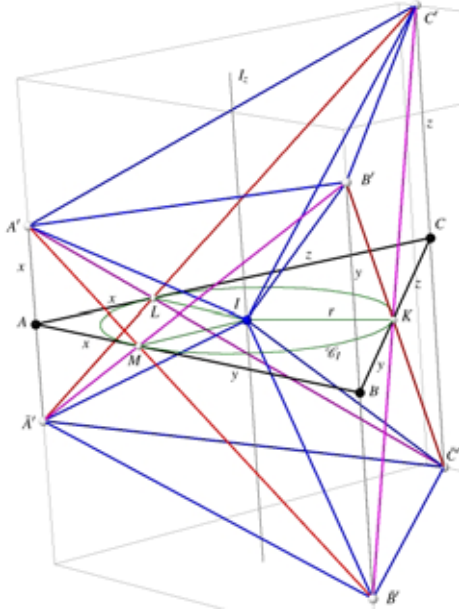


Figure 1: The two reguli: $A'B'$ and $\bar{A}'B'$.

Proof.

The simplest proof for the orthocentricity of the tetrahedron is the fact that the sum of the square of the length of a pair of opposite edges is a constant. For this proof (in [4], fig. 1), we used Pythagora's Theorem twice and we easily got $IA'^2 = AA'^2 + AM^2 + IM^2 = 2x^2 + r^2$ because the triangle AMA' is right and *isosceles*. It implies that the slope of $A'M$ in the vertical plane $A'AB$ is $\frac{\pi}{4}$. We use the same property to get $C'B'^2 = 2y^2 + 2z^2$ which yields the result . It was noted in [4, §3 - corollary 3.2] that the same proof for the symmetric tetrahedron also implied that the triplets $A'M\bar{B}'$ and similar are collinear. We conclude that we get the straight line supporting $B'K\bar{C}'$ from the one supporting $A'M\bar{B}'$ by a rotation about a vertical axis I_z going through the incenter and similarly for the straight line supporting $C'L\bar{A}'$. This proves that the nine vertices lie on a one sheet hyperboloid of revolution, equi-

lateral because the slope is $\frac{\pi}{4}$, which has I for center of symmetry and \mathcal{C}_I for throat circle.

Clearly the three segments $A'L$, $B'M$ and $C'K$ lie on one of the set of straight lines generating the hyperboloid while the other regulus is the set to which belong $A'M$, $B'K$ and $C'L$.

■

The way we defined the hyperboloid is linked only to the incircle \mathcal{C}_I and does not depend in any way of the choice of the other components of the triangle. Note that if \mathcal{C}_I is fixed, the choice of A, B, C is not totally arbitrary, we can choose the first vertex anywhere outside \mathcal{C}_I and then a second one has to be chosen on one of the half-lines (not going through A) supported by the tangents to the incircle going through A . The third one is then constrained. But any point of the plane outside the incircle can be the vertex of a triangle tangent to \mathcal{C}_I . Therefore, we have proved that when a triangle has \mathcal{C}_I for incircle, its Pavillet tetrahedron has the vertices of its upper triangle on a equilateral hyperboloid having \mathcal{C}_I for throat circle.

In [4] we called the sphere having the incircle as great circle the *incircle sphere*, its equation relative to the incenter is:

$$\mathcal{S}_I : x^2 + y^2 + z^2 = r^2. \quad (1)$$

We are going to see that our set of two symmetric tetrahedra has properties very similar relative to the hyperboloid as one of them relative to the incircle sphere. Hence we call this hyperboloid the *incircle hyperboloid*; its equation with the same origin is:

$$\mathcal{H}_I : x^2 + y^2 - z^2 = r^2. \quad (2)$$

3 The tangent planes.

Theorem 3.1 *The tangent plane at any vertex of one of the upper triangle goes through the orthocenter of the conjugate tetrahedron.*

Proof.

We recall that the segments $A'K$, $B'L$ and $C'M$ are three altitudes of the orthocentric tetrahedron [4, §1]. Their intersection, H'' , is such that its projection on the plane of the base triangle is G_e , the Gergonne point of the base triangle. This is valid for the conjugate tetrahedron. Now the

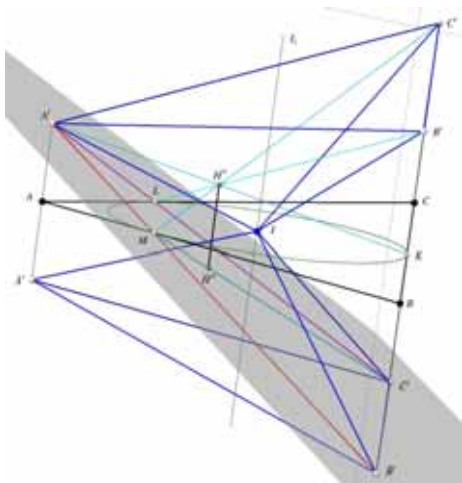


Figure 2: *The tangent plane at A' .*

tangent plane to the hyperboloid at one of these vertices is the plane define by both generatrices going through this vertex, e.g. the tangent plane at A' (fig. 2) is defined by the two lines $A'M$ and $A'L$ or, also, by $A'B'$ and $A'C'$ and therefore the line $C'M$ lies on the tangent plane at A' but $C'M$ being an altitude of the conjugate tetrahedron, H'' lies on the tangent plane to the hyperboloid at A' .

■

The intersection of the hyperboloid by an upper plane is a conic and consequently the three tangent planes at the upper vertices of one of the tetrahedron belong to the envelope of the tangent cone to the hyperboloid through the orthocenter. Therefore,

Theorem 3.2 *The pole of the upper plane of one tetrahedron about the hyperboloid is the orthocenter of its conjugate.*

Remark 3.1 *Note that when the intersection of one of the upper plane with the hyperboloid is a parabola, the angle between the upper plane and the horizontal plane is $\frac{\pi}{4}$ which from [6] and [4, §7] is known to be the special case when the outer Soddy circle of ABC degenerates to a line (see also [1, p. 62]).*

It is easy to check that the line of intersection of any two of the six tangent planes is either a generatrix of the hyperboloid, an altitude of one of the tetrahedron (excluding the one from the apex) or a line of support of the contact triangle.

3.1 Six more points.

We recall that both upper planes intersect the base plane along the Gergonne line (Γ , fig. 3) of the base triangle [4, §1.3]. As in [2], we call the points of intersections of the sides of the base triangle with the sides of the contact triangle K_g, L_g, M_g , the Nobbs points; they are collinear and lie on the Gergonne line of the triangle. We notice that the tangent plane at one vertex intersects the horizontal plane along a line which supports the contact triangle, e.g. the tangent plane at A' intersects the base plane along LM and therefore intersects the Gergonne line at one of the Nobbs points, K_g such that $(B, C, K, K_g) = -1$.

Remark 3.2 *The base triangle has to be scalene for the general case; if it is isosceles, e.g. at C , then M is the midpoint of AB and therefore M_g , its harmonic conjugate is at infinity. If the*

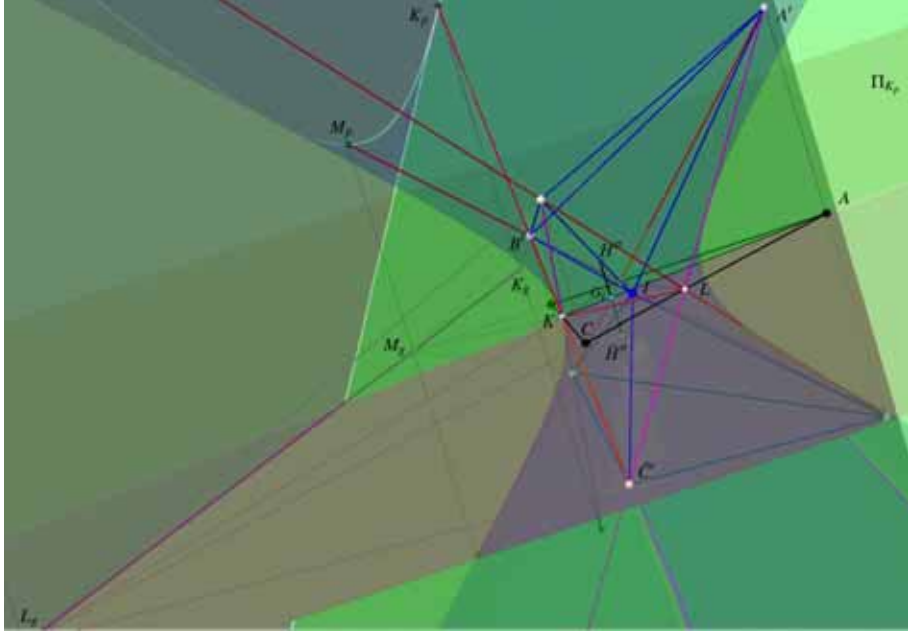


Figure 4: The plane Π_{K_p} tangent to the hyperboloid at K_p intersects the horizontal plane along AK

$(\Pi_{B'}, \Pi_{\tilde{C}'}, \Pi_K, \Pi_{K_p})$ on the horizontal plane.

4 Poles and polar properties.

In [4, §5], we proved that the incircle sphere, the sphere having the incircle as great circle, was a polar sphere of the orthocentric group formed by the vertices and orthocenter of the tetrahedron $IA'B'C'$. It implies that the orthocenter is the pole of the plane of the upper triangle with respect to this sphere. Here we have similarly proved Theorem 3.2.

It is shown in [4, §5, Theorem 5.4] that the line of orthocenters, $G_c H'' = \tilde{H}'' H''$, is the conjugate of the Gergonne line about the incircle sphere but this is also true for the hyperboloid because the conjugate of $H'' \tilde{H}''$ is the line of intersection of

two polar planes of points lying on this line, e.g. H'' and \tilde{H}'' , i.e. the two upper planes which, by symmetry intersect along the Gergonne line.

Similarly, if we consider the two altitudes going through the apex I and therefore through the center of both quadrics $(IH'', I\tilde{H}'')$, the conjugate line has to be a line at infinity. It is the line at infinity of the corresponding upper plane if we consider the sphere, the line at infinity of the conjugate upper plane if we consider the hyperboloid.

Going back to the five points I, A', B', C', H'' , it is an orthocentric group therefore the orthocentric tetrahedron $H''A'B'C'$ is self polar with respect to the incircle sphere. We compare with the polar tetrahedron of $H''A'B'C'$ with respect to the incircle hyperboloid. The polar plane of one of the orthocenters, e.g. H'' , is the upper plane of the conjugate tetrahedron defined by $\tilde{A}', \tilde{B}', \tilde{C}'$ while the polar plane of a vertex of the upper trian-

gle, A', B', C' , is the tangent plane to the hyperboloid at this point so that the polar tetrahedron is formed by one secant and three tangent planes.

Now the tangent plane at A' is also defined by the three points $\bar{H}'', \bar{B}', \bar{C}'$ (Theorem 3.1, fig. 2) so that the polar tetrahedron of $H''A'B'C'$ with respect to the hyperboloid is its conjugate $\bar{H}''\bar{A}'\bar{B}'\bar{C}'$.

Theorem 4.1 *The tetrahedra $H''A'B'C'$ and $\bar{H}''\bar{A}'\bar{B}'\bar{C}'$ are self polar relative to the incircle sphere but polar reciprocal with respect to the incircle hyperboloid.*

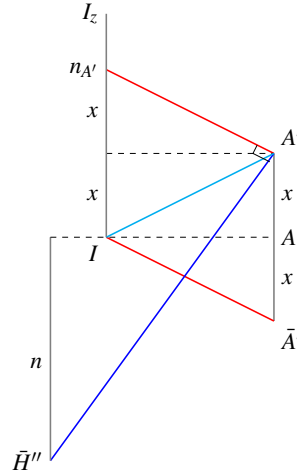
Therefore, the polarity about the incircle sphere keeps apart the elements of both tetrahedra while the polarity about the hyperboloid swaps them. It justifies the name of conjugate tetrahedra.

4.1 Normals to the hyperboloid at the fifteen points

We can also use the orthocentric group to find the normal to the hyperboloid at the vertices. As mentioned, the tetrahedron $\bar{H}''\bar{A}'\bar{B}'\bar{C}'$ is orthocentric and I is its orthocenter. Therefore $I\bar{A}'$ is orthogonal to the plane $\bar{H}''\bar{B}'\bar{C}'$ but this plane is the tangent plane to the hyperboloid at A' (fig. 2).

Therefore we have proved that the normal to the hyperboloid at one of the vertices of the tetrahedra is the parallel to the edge of the conjugate tetrahedra joining the apex to the corresponding conjugate vertex.

These normal will intersect the axis of the hyperboloid, the vertical line through the incenter I_z (fig. 2), because it is a surface of revolution. As shown on the following projection on the vertical plane IAA' , the altitude of the point of intersection is twice the tangential distance of the vertex of the base triangle to the incircle ($x = AL = AM$) and $\bar{A}'A' = In_{A'} = 2x$.



To find the normal to the upper and lower Nobbs point, e.g. K_p , we can use the fact that the relationship between the normal to the hyperboloid on a generator (here B', \bar{C}', K, K_p) and the corresponding point is algebraic and one to one. Then the cross ratio is preserved and because the set (B', \bar{C}', K, K_p) is harmonic then the intersection of the normals at these points with the axis of the hyperboloid is also harmonic. The normals at B' and \bar{C}' are known from above and intersect the axis at $n_{B'}$ and $n_{\bar{C}'}$. The tangent plane at K being vertical, the normal is the line KI which intersects the axis at I_z at I , therefore the normal at K_p intersects the axis at n_{K_p} harmonic conjugate of I about $n_{B'}$ and $n_{\bar{C}'}$. Hence, we can either construct n_{K_p} geometrically or compute its altitude z_{nk_p} with the equation

$$(z_{nk_p}, 0, 2y, -2z) = -1 \Rightarrow z_{nk_p} = \frac{4yz}{z-y}.$$

All other values can be deduced by simultaneous permutations and we have now all the values for the fifteen points found on the incircle hyperboloid. As seen above (Remark 3.2), if the triangle is isosceles, e.g. at C , $x = y$ then the corresponding point is at infinity.

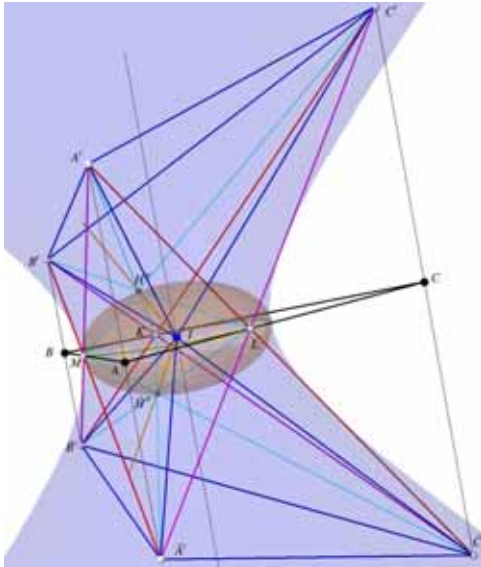


Figure 5: *The orthocenters lying on the ellipsoid inscribed in the throat circle of the hyperboloid.*

5 The incircle ellipsoid.

We have seen that the orthocenters of the conjugate tetrahedra play an important role in this geometric figure and so it becomes interesting to examine the locus of the orthocenters, also a surface, when $A', B', C', \bar{A}', \dots$ describe the hyperboloid.

This locus is easy to find because the orthogonal projection of the orthocenters on the plane of the base triangle is the Gergonne point of this triangle (G_e) and there is an invariant relation of the tetrahedron proved in [4, §6]: $g^2 + 3n^2 = r^2$ where r is the in-radius, g is the distance IG_e and n the altitude of the orthocenter about the plane of the base triangle $n = G_e H''$. Interpreting this relation in cylindrical coordinates about the in-center shows that this is the equation of an oblate ellipsoid of revolution

$$\mathcal{E}_I : x^2 + y^2 + 3z^2 = r^2, \quad (3)$$

which, again, is inscribed in \mathcal{E}_I , the throat circle of the incircle hyperboloid \mathcal{H}_I (2).

Therefore, because, to any set of points A, B, C forming a triangle circumscribed about \mathcal{E}_I correspond a single Gergonne point G_e and because, when A, B, C describe the outer part of the incircle, G_e describes the inner part, we see that, given a fixed circle, all the triangles which have this circle as incircle have their conjugate Paviilet tetrahedra attached to these two tangent and complementary surfaces (fig. 5).

5.1 Tangent plane to the ellipsoid at the orthocenters.

Because we know the surface on which the orthocenters lie, we may look for the tangent planes to the ellipsoid at these two points. As usual we have more than one way to find it, let's use a pencil of quadric.

- The incircle sphere and incircle hyperboloid form a pencil of quadrics. Their base curve is the incircle, counted twice, and at each point of the incircle they are tangent.
- Hence the incircle ellipsoid belongs to this pencil. Of course, we could as well check that, using (1), (2) and (3), we have

$$2\mathcal{S}_I - \mathcal{H}_I = \mathcal{E}_I.$$

- The polar planes of a point with respect to a pencil of quadrics form a pencil of planes, we apply this to one of the orthocenters.
- The polar plane of an orthocenter about the incircle sphere is the corresponding upper plane [4, Theorem 5.2].
- The polar plane of this orthocenter about the incircle hyperboloid is the conjugate upper plane (cf. Theorem 3.2).
- Both planes intersect at the Gergonne line of the base triangle (§3.1, fig. 6).

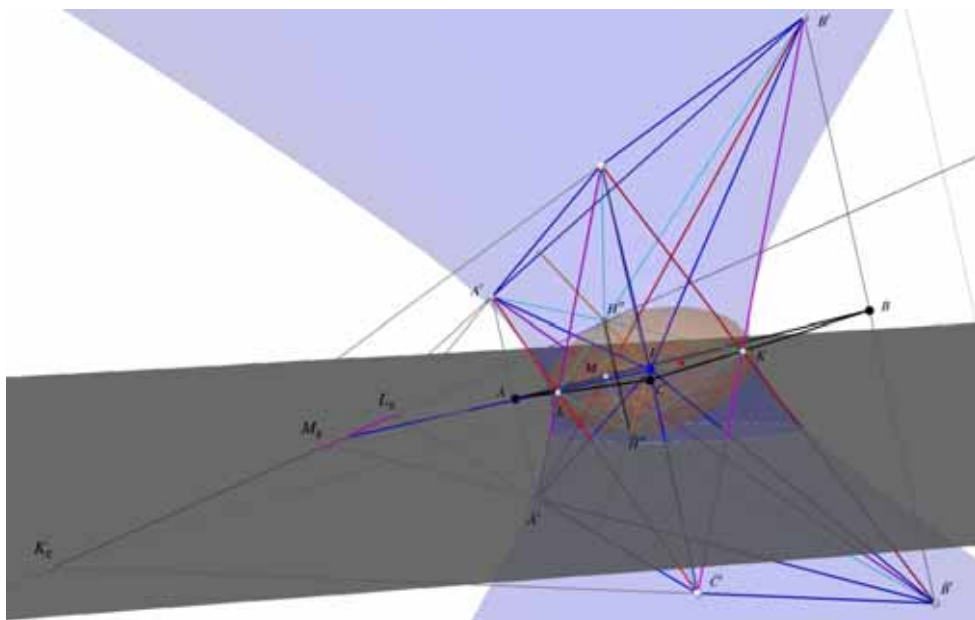


Figure 6: *The tangent planes to the ellipsoid at the orthocenters.*

Giving one more constraint to the triangle would replace this two parameters problem by a one parameter problem; the vertices and the orthocenters would describe a curve which would lie on one of these two surfaces. As a simple example, we can consider the degenerate case of the Soddy outer circle (Remark 3.1). In that case the orthocenter will be such that the distance $IG_e = \frac{r}{2}$, and therefore the orthocenters will describe two symmetric horizontal circles on the ellipsoid. This example and others deserve further studies. It was mentioned in [4] that Richard Guy had extended the properties of the Pavillet tetrahedron to the excircles [1, §5.3] so that another possible development are the *excircle hyperboloids*, the extension should be straightforward but the relationships between the four hyperboloids can reveal interesting properties.

Bibliography

- [1] R. K. Guy. The triangle. on line, 2011. URL <https://xa.yimg.com/kq/groups/386540/1852736325/name/quadrations.pdf>.
- [2] A. Oldknow. The Euler-Gergonne-Soddy triangle of a triangle. *The American Mathematical Monthly*, 103: 319–329, April 1996.
- [3] A. Pavillet. The orthocentric tetrahedron of a triangle, new properties and inverse problem. In *Proceedings of the 15 th International Conference on Geometry and Graphics (ICGG 2012)*, pages 569–578. 08 2012. ISBN 978-0-7717-0717-9. URL http://people.sutd.edu.sg/~axel_pavillet/?page_id=100.
- [4] A. Pavillet. The orthocentric tetrahedron

of a triangle. *Forum Geometricorum*, accepted 2012. URL http://people.sutd.edu.sg/~axel_pavillet/wp-content/uploads/2013/07/Tetrahedron-1-Final-PDF.pdf.

- [5] A. Pavillet and V. V. Shelomovskii. Elementary proof of pavillet tetrahedron properties. *Research Journal of Mathematics & Technology*, 1: 87–96, October 2012. URL <https://php.radford.edu/~ejmt/ContentIndex.php>.
- [6] A. Vandeghen. Soddy’s circles and the De Longchamps point of triangle. *The American Mathematical Monthly*, 1964.
- [7] G. Weiss. Is advanced elementary geometry on the way to regain scientific terrain? In *Proceedings of the 15 th International Conference on Geometry and Graphics (ICGG 2012)*, pages 793–804. 08 2012. ISBN 978-0-7717-0717-9.

He has published numerous articles in the scientific and technological fields. He is knight of the National Order of Merit from France and was awarded the Meritorious Service Medal from the United States.

About the author

Axel Pavillet is a graduate of Ecole Polytechnique in Paris and has a M. Eng in industrial engineering from Ecole Nationale Supérieure des Techniques Avancées. He spent most of his career as a professional engineer and a high level executive for the French Government. He worked in France, the United States, Argentina and Canada. In 2000, he went back to academia to earn a M.Sc. in Computer Science and a Ph.D. in Mathematics from Université du Québec à Montréal in 2001 and 2004, respectively. Since then, he dedicates his time to teaching and research. He is presently teaching at Singapore University of Technology and Design (SUTD) and a licensed P. Eng. registered with the Association of Professional Engineers and Geoscientists of Alberta.

INTERACTION AMONG COURSES

Jelena BEBAN-BRKIĆ ¹ and Marija ŠIMIĆ HORVATH ²

¹Faculty of Geodesy, University of Zagreb, Croatia

²Faculty of Architecture, University of Zagreb, Croatia

ABSTRACT: In our daily contact with students within different mathematics/geometry courses we experience that they are often unable to use and link the knowledge they have acquired. In order to overcome such situations we believe that there should be linkage between courses. The interaction among courses may be realized in several ways: cooperation between lecturers with the goal of expanding the mathematical and lecture discourse of a particular class, preparing "joint" posters, and giving "joint" lectures.

Keywords: mathematics education, interaction, spatial ability

1. INTRODUCTION

Do you find that in some mathematics/geometry course students are unable to use and link the knowledge they have acquired in some other course?

The described situation occurs frequently enough in our everyday teaching that we should not speak of it as an accident or a local phenomenon. We should also not say that it is due to newly arisen circumstances caused by the partially altered and reduced curricula in general subjects at technical faculties that followed the adoption of the Bologna Process.

2. BACKGROUND

The problem, insofar as teachers experience it, has existed for a rather long time. Perhaps it was not that prominent before because the abundant classes for mathematics/geometry courses provided room for a deeper and broader elaboration of teaching content. So students, by themselves, were able to make a link with the knowledge they had acquired.

In our opinion, the first warning signs appeared as early as 1990s. During those years in the

instruction of mathematics, the oral exams were gradually abolished while geometry courses were frequently removed and/or transformed into computer graphics or computer-aided design (CAD).

The consequence of abolishing the oral examinations is the predomination of the practical over the theoretical part of the material, which is negatively reflected in understanding and linking it to what has been learned. Interventions in geometry courses do not provide the teachers enough space to stimulate to an appropriate extent the development of spatial ability among the majority of students. Therefore, it is more difficult to encourage them to think spatially.

Further decreasing teaching hours (Bologna Process) and segmented examinations also distance students (through no fault of their own) from linking it to knowledge acquired in various other courses.

3. LOOKING AHEAD

The perception of this situation as well as the similar one presented in [2] have impelled us to think about how and what to do regarding

this question within the existing schedule of mathematics/geometry subjects ([5], [6]), while not intervening in, though enhancing their learning outcomes.

We are aware that the mathematics education of students in the engineering disciplines must be laid carefully and that this takes time ([3]). Still, we believe that within time and knowledge that we have at our disposal we could mediate in an appropriate way in linkage between courses. The linkage/interaction among courses may be realized in several ways: cooperation between lecturers with the goal of expanding the mathematical and lecture discourse of a particular class, preparing "joint" posters, and giving "joint" lectures.

Hereinafter are given parts of such a poster.

4. POSTER EXAMPLE

4.1 Coordinate System

Points P in 3-space (three dimensional space) are represented by ordered triples of real numbers (p_1, p_2, p_3) ; the numbers p_1, p_2 , and p_3 are called the *Cartesian coordinates* of P . To construct such a representation, we choose three mutually perpendicular lines that meet at a point O (called the *origin*). These lines are called *x-axis*, *y-axis*, and *z-axis*.

Rectangular coordinate systems in 3-space fall in two categories, *left-handed* and *right-handed*. Within the *descriptive geometry* course we use the left-handed coordinate system, while in *analytical geometry* the right-handed is used.

At this point we suppose that students are familiar with the basic knowledge of analytical geometry, linear algebra and descriptive geometry/orthogonal projection ([1], [4]).

So, our idea is not to start from the beginning but to collect and relate knowledge students already have acquainted.

4.2 A Point

Approach 1

The method of orthogonal projection consists

of projecting an object on three mutually perpendicular planes by rays perpendicular to these planes. One plane is called the *horizontal* plane of projection and it is denoted by Π_1 , the another one is called *vertical* denoted by Π_2 , and the third one is called the *profile* plane of projection denoted by Π_3 . The planes are infinite, unbounded and not transparent. The lines of intersection of these planes are called the coordinate axes. Hence,

$$\begin{aligned}\Pi_1 \cap \Pi_2 &\equiv x, \\ \Pi_1 \cap \Pi_3 &\equiv y, \\ \Pi_2 \cap \Pi_3 &\equiv z.\end{aligned}$$

A projection drawing on which the projection planes are brought coincidence (in specific manner) is called an *orthographic representation* (see Figure 2). The negative semiaxes are not indicated in the orthographic representation.

Let the point T be given by (x, y, z) .

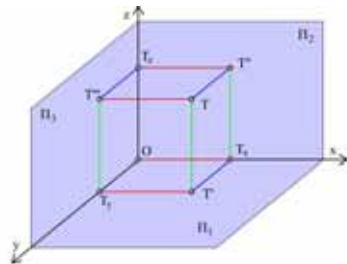


Figure 1: Coordinate System (the left-handed)

x defines the distance of the point T to the profile plane Π_3 , y to the vertical plane Π_2 , and z to the horizontal plane Π_1 .

T' stands for the orthogonal projection of the point T to Π_1 , T'' for its orthogonal projection to Π_2 , and T''' for its orthogonal projection to Π_3 (see Figure 1).

The orthographic representation of the situation presented in Figure 1 is given by Figure 2.

T' is called a *top view* of the point T , T'' is called the *front view* of T , and T''' is called the *side*

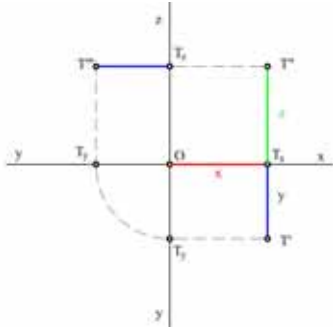


Figure 2: Orthographic Representation

view.

The orthogonal projections of an object are represented only on two planes Π_1 and Π_2 . In this case there is only x -axis labelled in the orthographic representation.

Approach 2

The Standard Basis Vectors

With each vector \vec{d} we associate the point (a_x, a_y, a_z) where \vec{d} terminates, and conversely, i. e. we have a relation

$$\text{point}(a_x, a_y, a_z) \Leftrightarrow \text{vector } \vec{d} = \{a_x, a_y, a_z\}$$

We call a_x, a_y , and a_z the *components* of \vec{d} , or when we think of \vec{d} as a point, its *coordinates*. The components a_x, a_y , and a_z of \vec{d} are the lengths of the projections of \vec{d} along the three coordinate axes.

A *position vector*, also known as *radius vector*, is a vector which represents the position of a point in space in relation to the origin O . It is usually denoted by \vec{r} . In this sense, for the point A ,

$$\vec{r} = \vec{OA}.$$

Vectors $\vec{i} = \{1, 0, 0\}$, $\vec{j} = \{0, 1, 0\}$, $\vec{k} = \{0, 0, 1\}$ are unit vectors along the three coordinate axes, as shown in Figure 3. We call them the *standard bases vectors*.

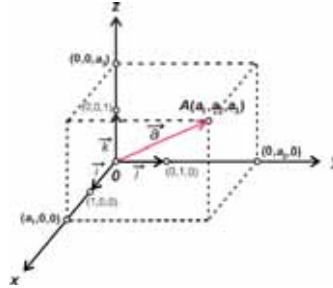


Figure 3: Coordinate System (the right-handed)

If \vec{d} has components a_x, a_y, a_z , then

$$\vec{d} = a_x \vec{i} + a_y \vec{j} + a_z \vec{k}.$$

If A has coordinates (a_x, a_y, a_z) , then

$$\vec{r} = \vec{OA} = a_x \vec{i} + a_y \vec{j} + a_z \vec{k}.$$

Sometimes a vector is positioned so that its initial point is not at the origin. If the vector $\vec{T_1T_2}$ has initial point $T_1(x_1, y_1, z_1)$ and terminal point $T_2(x_2, y_2, z_2)$, then $\vec{T_1T_2}$ is the difference of the corresponding position vectors $\vec{OT_2}$ and $\vec{OT_1}$ (see figure). So,

$$\begin{aligned} \vec{T_1T_2} &= \vec{OT_2} - \vec{OT_1} = \{x_2, y_2, z_2\} - \{x_1, y_1, z_1\} = \\ &= \{x_2 - x_1, y_2 - y_1, z_2 - z_1\}. \end{aligned}$$

4.3 A Line

Approach 1

Every line is uniquely determined by two points. If the point T is incident with the line p then the projections of the point T must be incident with like projections of p . The opposite direction does not always hold. Representing the line by means of the orthogonal projection, two important points specify the line, the horizontal and the vertical trace point.

The trace of the line is the point of intersection of the line with a projection plane.

The horizontal trace P_1 is the point of intersection of the line p with the horizontal plane Π_1 .

The vertical trace P_2 is the point of intersection

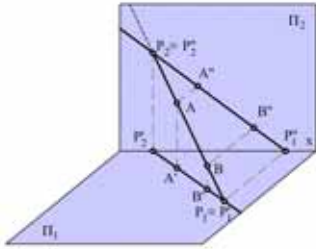


Figure 4: Line in 3-Space

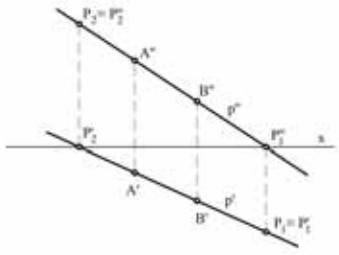


Figure 5: Line in Orthographic Representation

of the line p with the vertical plane Π_2 .

Hence, the trace belonging to one projection plane coincides with its like projection, $P_1 \equiv P'_1$ and $P_2 \equiv P'_2$.

The line p in Figure 4 is given by points A and B .

Description of Figure 5:

$p' \equiv A'B'$ - the top view of the line p ,

$p'' \equiv A''B''$ - the front view of the line p ,

$P'_1 = p' \cap x \quad \wedge \quad P_1 \equiv P'_1$,

$P'_2 = p'' \cap x \quad \wedge \quad P_2 \equiv P'_2$.

Approach 2

Any line, whether in 2-space or in 3-space, can be described by giving two pieces of information:

- a single point on the line (called the *initial point*), and
- a vector giving the direction of the line (called a *direction vector*).

The initial point and direction vector of a line are not unique:

- the initial point can be any point on the line,
- the direction vector can be any non-zero vector that points in the appropriate direction.

Let the initial point $T_0(x_0, y_0, z_0)$ and the direction vector $\vec{p} = \{l, m, n\}$ of the line p be given.

We will find the equation for p by using *vector*

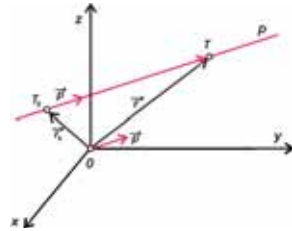


Figure 6: A line parallel to a given vector

addition and scalar multiplication. Let $T(x, y, z)$ be any point on p .

The corresponding radius vectors are:

$\vec{OT}_0 = \vec{r}_0 = \{x_0, y_0, z_0\}$, $\vec{OT} = \vec{r} = \{x, y, z\}$.

For any point on p ,

$$T \in p \Rightarrow \vec{T_0T} = \lambda \vec{p}, \quad \lambda \in R,$$

$$\vec{T_0T} = \vec{r} - \vec{r}_0$$

$$\vec{r} = \vec{r}_0 + \lambda \vec{p} \quad (1)$$

Vector form of the equation of a line

(1) in terms of components can be written as $x\vec{i} + y\vec{j} + z\vec{k} = (x_0\vec{i} + y_0\vec{j} + z_0\vec{k}) + \lambda(l\vec{i} + m\vec{j} + n\vec{k})$ from which it follows

$$\begin{aligned} x &= x_0 + \lambda l \\ y &= y_0 + \lambda m \\ z &= z_0 + \lambda n \end{aligned} \quad (2)$$

Parametric equation of a line

Eliminating λ in (2) one gets

$$\frac{x-x_0}{l} = \frac{y-y_0}{m} = \frac{z-z_0}{l} (= \lambda) \quad (3)$$

Canonical/standard equation of a line

4.4 A Plane

Approach 1

The plane P in Figure 7 is in relative position

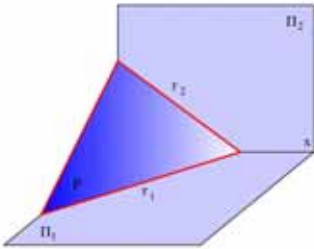


Figure 7: Plane in 3-Space

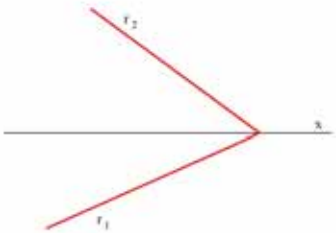


Figure 8: Plane in Orthographic Representation

with respect to the projection planes Π_1 and Π_2 . The lines of intersection between the plane P and the projection planes are called the *traces* of P and denoted by r_1 and r_2 .

In an orthographic representation (see Figure 8) the plane is uniquely determined by its traces, the

horizontal trace and the *vertical trace*. Hence,

$$r_1 = P \cap \Pi_1, r_1 - \text{the horizontal trace of } P,$$

$$r_2 = P \cap \Pi_2, r_2 - \text{the vertical trace of } P.$$

They have a common point on the x -axis.

Figures 9 and 10 present the planes that are perpendicular to one of the projection planes, the *horizontal projecting plane* $\Sigma(s_1, s_2)$ and the *vertical projecting plane* $\Delta(d_1, d_2)$.

If a figure lies in a plane perpendicular to some

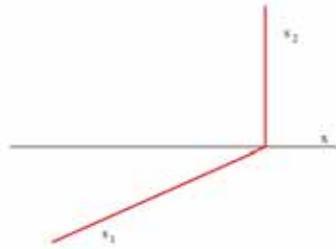


Figure 9: $\Sigma(s_1, s_2) \perp \Pi_1, s_2 \perp x$

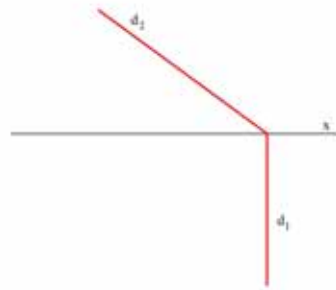


Figure 10: $\Delta(d_1, d_2) \perp \Pi_2, d_1 \perp x$

plane of projection, its projection onto this plane coincides with the like trace of projecting plane.

Approach 2

In analytical geometry a plane in 3-space can be

specified by giving its inclination and one of its points. The inclination of a plane is provided by a nonzero vector, called a *normal*, which is perpendicular to the plane.

The following approach to finding the equation of a plane assumes knowledge of the *scalar product*.

Let the point $T_0(x_0, y_0, z_0)$ and the nonzero vec-

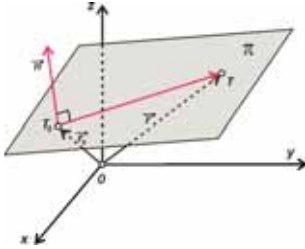


Figure 11: Plane with normal vector

tor $\vec{n} = \{A, B, C\}$ as a normal of the plane π be given (see Figure 11). The plane consists of those points $T(x, y, z)$ for which the vector $\overrightarrow{T_0T}$ is orthogonal to \vec{n} , that is,

$$\vec{n} \cdot \overrightarrow{T_0T} = 0 \quad (4)$$

Denoting $\vec{r_0} = \overrightarrow{OT_0} = \{x_0, y_0, z_0\}$ and $\vec{r} = \overrightarrow{OT} = \{x, y, z\}$, formula (4) can be rewritten as

$$\vec{n} \cdot (\vec{r} - \vec{r_0}) = 0 \quad (5)$$

Vector form of the equation of a plane

(5) in terms of components can be written as $\{A, B, C\} \cdot \{x - x_0, y - y_0, z - z_0\} = 0$, or

$$A(x - x_0) + B(y - y_0) + C(z - z_0) = 0$$

Algebraic/point-normal form of the equation of a plane

(6)

By multiplying out, collecting terms, and denoting $D = -(Ax_0 + By_0 + Cz_0)$, (6) can be rewritten in the form

$$Ax + By + Cz + D = 0$$

Canonical/general form of the equation of a plane

(7)

If $D = 0$ it is a matter of a plane through the origin.

5. CONCLUSIONS

This joint work is an example of the interaction among courses. A few tasks from geometry courses and from analytic geometry and linear algebra courses are displayed as a poster. At the same time, we are showing different approaches: the spatial draft of the task, the task solved using the orthogonal projection method, and the algebraic approach to the task.

We consider such posters to be educational material that could be helpful to students in our courses in connecting with the knowledge they have acquired.

REFERENCES

- [1] H. Anton and C. Rorres. *Elementary Linear Algebra*. John Wiley and Sons, Inc., New York, 2000.
- [2] J. Beban-Brkić and D. Jovičić. Distance From Point to Line and From Point to Plane. *MIŠ*, 42: 61–68, 2007.
- [3] B. Divjak et al. *Learning Outcomes in Higher Education*. TIVA Varaždin, Faculty of Organization and Informatics, Varaždin, Croatia, 2008.
- [4] N. Krylov, P. Lobandiyevsky, and S. Men. *Descriptive Geometry*. Mir Publishers, Moscow, 1974.
- [5] <http://e-ucenje.geof.unizg.hr/>
- [6] <http://virtual.arhitekt.hr/>

ABOUT THE AUTHORS

1. Jelena Beban-Brkić, Ph.D. math., is an assistant professor at the Department of Mathematics and Physics, Faculty of Geodesy, University of Zagreb. She performs a number of duties at the Faculty, ECTS Coordinator, President of the Commission for

textbooks and scripts, et cetera. Her main research interests are in geometry/non-Euclidean geometry and mathematical education. She can be reached by e-mail: jbeban@geof.hr

2. Marija Šimić Horvath, Ph.D. math., is an assistant professor at the Department of Mathematics, Descriptive Geometry and Perspective, Faculty of Architecture, University of Zagreb. Her research interests are in non-Euclidean geometry and mathematical education. She is actual president of Croatian Society for Geometry and Graphics. She can be reached by e-mail: marija.simic@arhitekt.hr

INTRODUCING 3D MODELING INTO GEOMETRY EDUCATION AT TWO TECHNICAL FACULTIES AT THE UNIVERSITY OF ZAGREB

Sonja GORJANC, Helena HALAS, Ema JURKIN
University of Zagreb, Croatia

ABSTRACT: In this paper we present how 3D computer modeling is introduced into the teaching of Descriptive geometry and Perspective courses at two technical faculties at the University of Zagreb.

Keywords: geometry education, 3D modeling, Rhinoceros, e-learning

1. INTRODUCTION

Teaching constructive geometry, mainly within the courses of descriptive geometry, has a long and well-established tradition at the technical faculties in Croatia. But, excluding the University of Rijeka, long-time insistence on the traditional way of teaching resulted in lagging behind most countries of European Union in the application of computer 3D CAD packages and implementation of educational content in e-learning systems at the beginning of 2012. During the year 2012 the authors, together with nine other teachers, worked on the project *Introducing 3D Modeling into Geometry Education at Technical Colleges* supported by the Fund for the Development of the University of Zagreb. The focus of the project was creating a basic repository of educational material related to common teaching topics and those customized to profiles of each faculty (<http://www.grad.hr/geomteh3d/>). The content of this project is described in detail in [2].

In this paper we present a part of the educational material, made within the aforementioned project, related to 3D computer modeling with the program *Rhinoceros* and our experience of using this material in teaching descriptive geometry courses at the Faculty of Civil Engineering (FCE) and the Faculty of Mining, Geology and Petroleum Engineering (FMGPE) through last two academic years.

2. DESCRIPTIVE GEOMETRY COURSES

In aforementioned faculties, teaching of Descriptive geometry is held in the first year of study: (2+3) in 1st semester at FCE, (2+2) in 1st and 2nd semester at FMGPE. For most of the students, this is their first encounter with this subject. The course contains two separate parts. In the first part, students work with a ruler and compasses and are introduced into the method of Monge projections. Constructive procedures are being explained in detail and students are consequently capable to construct orthogonal projections of simple geometric objects (prisms, pyramids, cones and cylinders) in general position to the planes of projections. After that, basics of axonometry are being taught and students draw in hand one simple object. Meanwhile, in exercise classes, they draw axonometric images of more complicated objects using *Rhinoceros*. Afterwards, all matter being taught in lectures is also constructively handled in this program.

Rhinoceros was chosen for a few reasons: a good experience of colleagues from the University of Innsbruck, a relatively low cost of the educational lab licence and the free trial versions available for students' downloads. Due to the lack of time, students are not taught basics of *Rhinoceros* in exercise classes. For that matter, the repository (that is made within the aforementioned project) contains more than 50 five-minute video clips in which constructions in

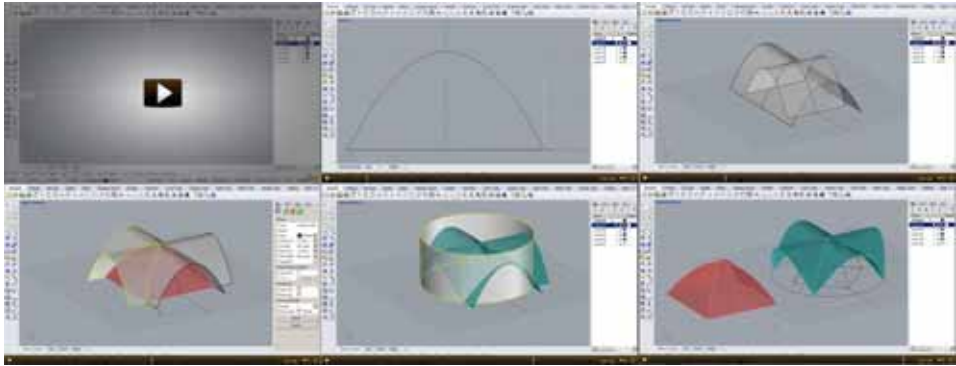


Figure 1: Screenshots from the video clip related to the intersection of parabolic cylinders.

Rhinoceros are being explained.

The video material was produced using free online program *Jing* which creates five-minute videos with sound. Using these video clips, students are introduced to the basics of the program *Rhinoceros* on their own, while they still draw in hand in class. For each of the later teaching unit, students are given a few solved representative examples so they can master the program by themselves. In this manner, we achieved the following: students work at home to learn the program and then come to the exercise classes and solve geometric problems concerning descriptive geometry.

All of the teaching material created is organised in teaching units and is available online on e-learning web pages. University of Zagreb developed this pages using e-learning platform Merlin, the system based on the learning management system Moodle. Our application of e-learning has level 2, but in some segments level 3, [1]. On these web pages students can also find information about their grades. We grade their exercises and mid-term exams. Due to the relatively large number of students (each teacher handles cca 50 students), we refused to use the option for uploading student exercises. That

would diminish our control over their work and the communication about their work would be more complicated. Because of that, we only grade exercises solved in class in front of the teacher, or optionally in teacher consultations. Each exercise class is graded.

Each student workgroup in computer lab has 10-12 students. This number allows the teacher to have good insight in the work of each student, gives the teacher the ability to examine acquired knowledge and deepen it if necessary, according to individual needs of every student. For each lesson, a student is given a few exercises (mostly 3) in advance. He/she prepares solutions to these exercises at home before class. If he/she successfully solves these exercises and answers a few questions regarding the material being taught, he/she is graded good (C). To receive a higher grade, the student must solve one or two additional exercises given in class by the teacher. Duration of the computer lab class is 60 minutes.

Topics managed in *Rhinoceros* are: solid modeling, cross sections of surfaces, intersection of surfaces and terrains with roads and layers. Some examples are shown in figures 2-5.

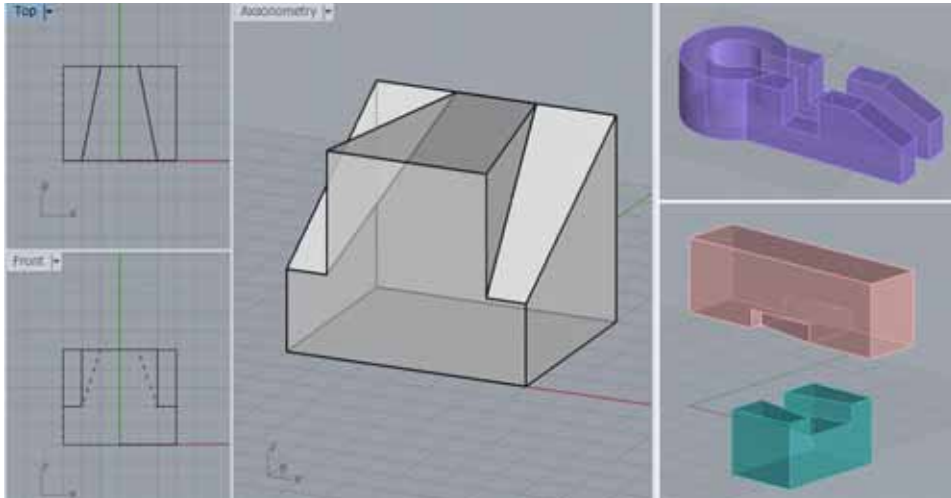


Figure 2: Three solved examples for solid modeling. Students are given two principal views of the object and their objective is to construct its axonometric image in *Rhinoceros*. This is the content of the first lesson in the computer lab.

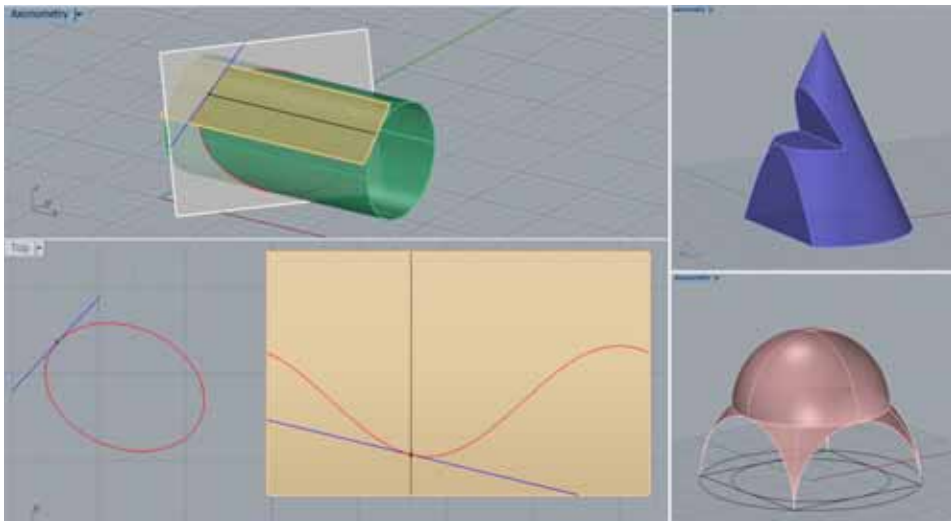


Figure 3: Solved examples for intersections of surfaces and planes. In this topic students handle cone, cylinder and sphere sections, tangent planes of these surfaces and tangents of intersection curves.

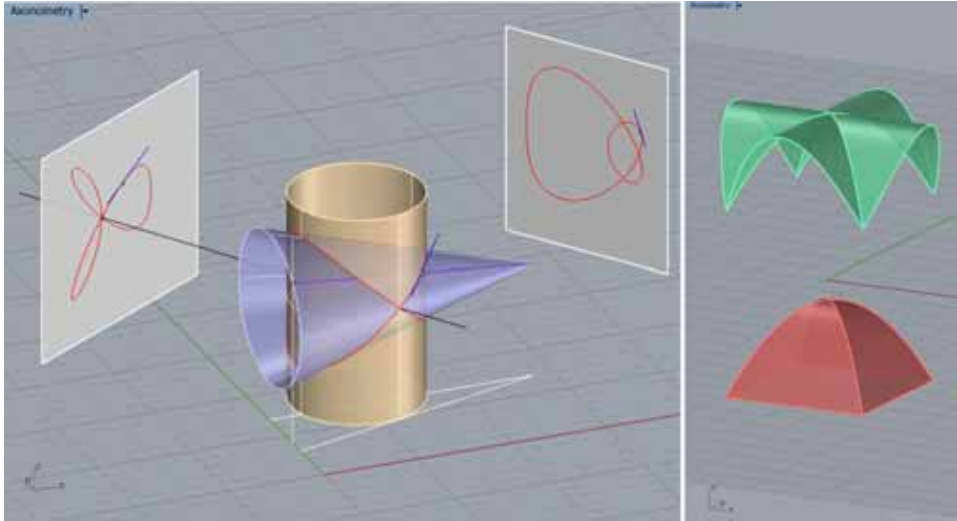


Figure 4: Two solved examples for intersections of surfaces. We deal with intersections of cones, cylinders and spheres, the tangents, plane projections and splitting of intersection curves.

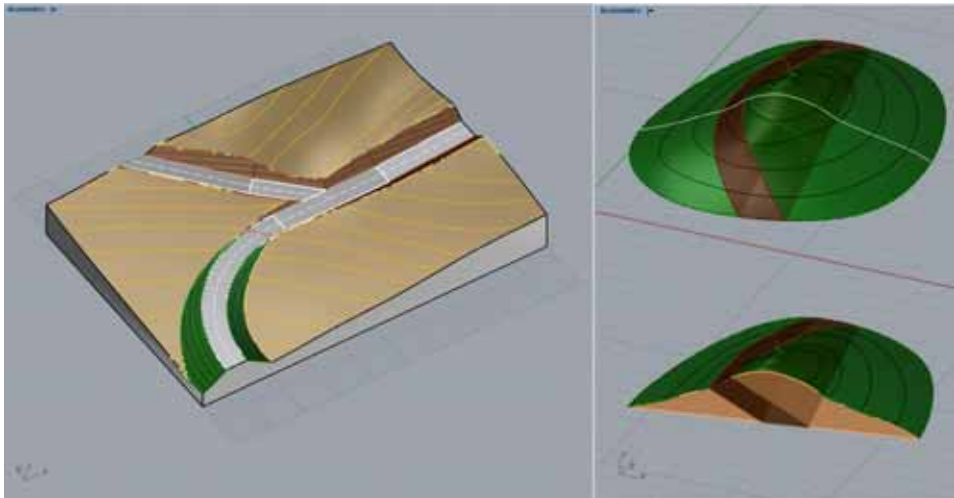


Figure 5: Examples of solved exercises for the situation of earthworks beside the road on the terrain and layers on the terrain. Before they model these exercises in *Rhinoceros*, students learn about the projection with elevations and solve the same exercises in hand in this projection method.

3. OPTIONAL GEOMETRICAL COURSE FOR MASTER STUDENTS

A decade ago, when Bologna Process caused educational reform in Croatia, geometry teachers of Department of Mathematics at FCE suggested two optional courses for master students: Basics of differential geometry and Perspective. Until now, students showed no interest in the first course - we believe this is because it contains ample amount of higher mathematics, while Perspective aroused certain attention. In period 2009-2012, each academic year 7-9 students enrolled in Perspective. The content of this course included the following: basics of perspective, construction of perspective images of rooftops, ruled surfaces, roads on the terrain. These constructions were made in hand or in some CAD program in 2D.

When we implemented computer 3D modeling in teaching first year courses in academic year 2012/2013, we did the same in Perspective. In that first year, students already showed greater interest in that course (12 students enrolled). Next academic year 2013/2014 that course had 34 students, mostly from Structural engineering and Transportation engineering programme.

Students who enroll in this course are in majority in their last semester of master programme during which they have never encountered 3D computer modeling before. That is the reason we can use our educational material prepared for the course Descriptive geometry in teaching Perspective. When generations which were taught 3D modeling in their first year of studying come to master level, our plan is to introduce *Grasshopper* as well as the basics of *Python* scripting for *Rhino* and *Grasshopper* (only on the informational level).

Perspective is being taught in the final semester of master programme, so by FCE's decree, whole class is held during 8 weeks, giving the students enough time to prepare their theses. Hence, work is very intense (8 classes per week) and workgroups consist of 8-9 students.

During the first two weeks, students are introduced to the basics of perspective drawing and are enabled to construct perspective images of simple geometric objects. In the third week, 3D modeling is introduced and they construct perspective images of objects using program *Rhinoceros*. During next two weeks, students are acquainted with quadric surfaces and ruled surfaces with emphasis on quadric ruled surfaces. Furthermore, examples of these surfaces in civil engineering are shown and they are acquainted with geometric interpretation and visualisation of notions of differential geometry of these surfaces (notions such as classification of points on a surface, principal curvatures, normal curvature, principal directions). In sixth week the topic is conoids of third and fourth order. During last two weeks, students model situations of earthwork beside roads on terrains.

This course is optional, held in the final year of studies and has very small study groups so the concept of exercises and homework assignments differs a lot from Descriptive geometry. In classes students solve exercises available online. Working version of the collection of exercises (in Croatian) is available on line <http://www.grad.hr/sgorjanc/perspektiva-vjezbe.pdf>. Homework assignments are graded. They are made by students at home and uploaded over system Merlin. For four of these assignments only the topic is given (solid, dome and vault, coverings with quadric ruled surfaces, coverings with higher degree conoids) so a student has the liberty to choose which object (something they know from real life or something imaginary) he/she will model. In the beginning of each week, we give an overview of all homework assignments from the week before. Students enjoy this very much because it inspires them and induces cooperation. The last assignment is to model given road on a terrain. This is done in class and immediately uploaded to Merlin. If there is something left to finish, the student can finish this assignment at home and upload it again. Some examples of student homework assignments are shown in figures 6 - 10.

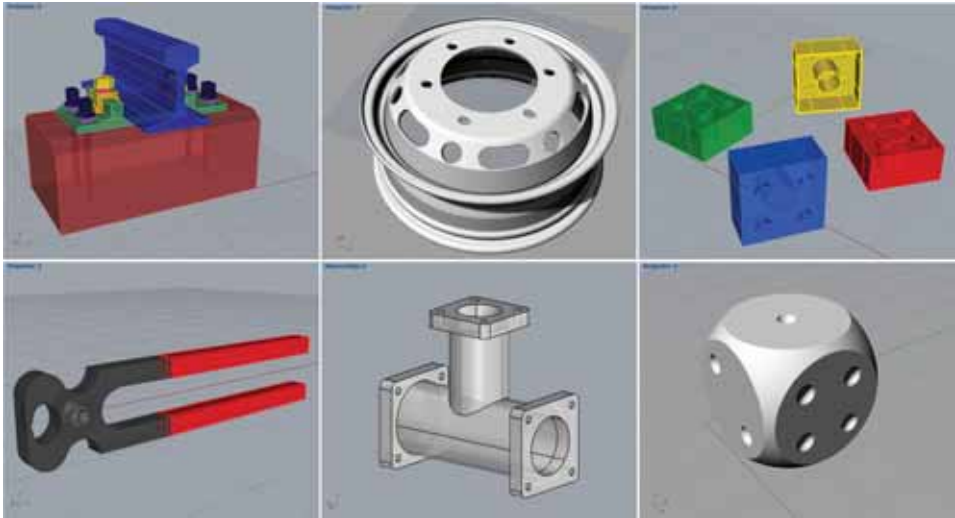


Figure 6: Student homework assignments on the subject *solids*.

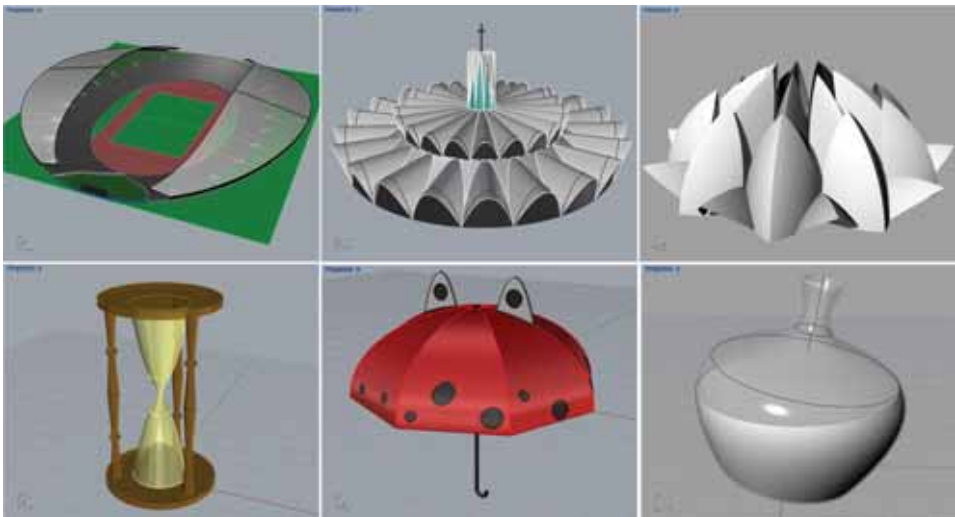


Figure 7: Student homework assignments on the subject *surfaces of revolution, domes and vaults*.

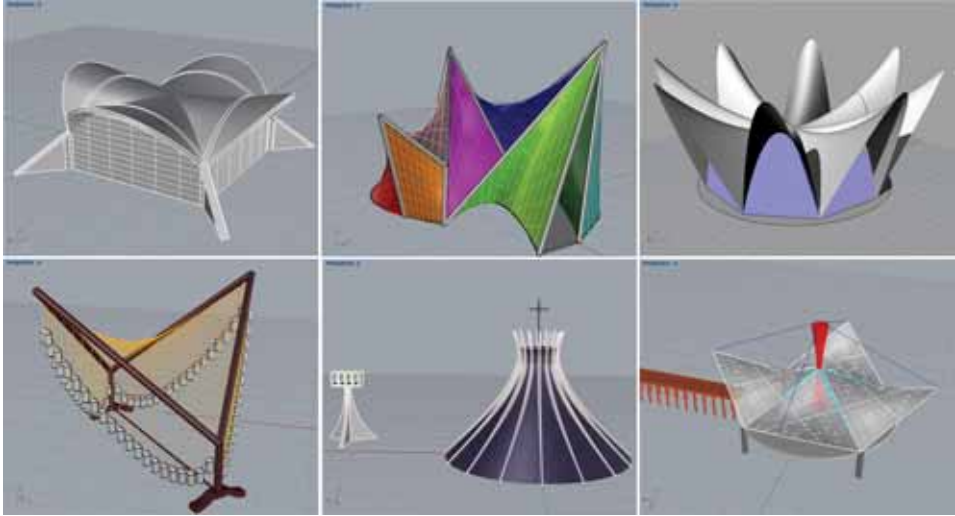


Figure 8: Student homework assignments on the subject *hyperbolic paraboloid and hyperboloid of one sheet*.

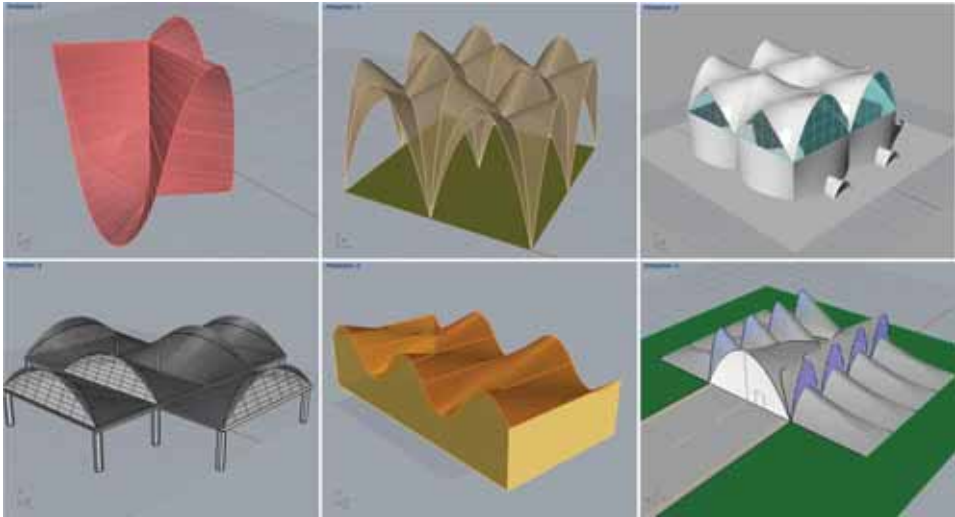


Figure 9: Student homework assignments on the subject *3rd and 4th order conoids*.

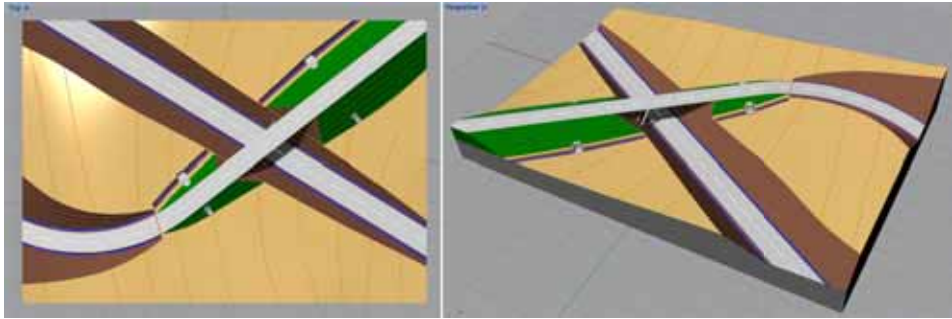


Figure 10: Two views from student homework assignment related to *earthworks beside the road on the terrain*, made by student Šime Bezina.

4. CONCLUSIONS

As stated in the Introduction, because of subjective reasons, University of Zagreb introduced 3D modeling in teaching geometry at the technical faculties much later than other European countries which traditionally nourish descriptive geometry as inevitable segment of education for engineers. However, foreign experience enabled our current teachers to master computer-aided geometry teaching very quickly and on a satisfactory level. We were able to use plenty of positive experience from our colleagues (mostly from Austria), and we could rely on advanced technology and methodology. Regarding all of that, we avoided many possible beginner's errors. Our greatest inspiration were thoughts stated in the last section of the article [3].

Here we would like to point out some of our positive experience:

- Resulting repository of educational material, made within the project mentioned in the Introduction, is of high quality due to the collaboration of teachers from four different technical faculties.
- Free access to the repository has had positive effect on a wider teaching community.
- Video material for mastering basics of the CAD program allows students to learn indepen-

dently at their own pace. On the other hand, it enables teachers, who create that material, to connect students' learning programming functions with their acquiring the knowledge of geometry.

- Homework assignments which are given only by the theme and give the student freedom to choose the object they would like to model are proven to be very inspiring and they evoke creativity within students. Such assignments make students search geometrical properties in their surroundings, and they can be used as a source of ideas for future exercises.

REFERENCES

- [1] J. Beban Brkić, S. Gorjanc, Ž. Tutek. New Challenge for Developers of E-infrastructure. *Proc. 15th ICGG Montreal*, 2012.
- [2] S. Gorjanc and E. Jurkin. Introducing 3D Modeling into Geometry Education at Technical Colleges. *The Visual Language of Technique: Volume 3 - Heritage and Expectations in Education*, editor L. Cocchiarella, Springer, 2014. (to appear)
- [3] H. Stachel. The status of today's Descriptive geometry related education (CAD/CG/DG) in Europe. *40th anniversary*

sary of the Japan Society for Graphics Science, Tokyo, May 12-13, 2007.

ABOUT THE AUTHORS

1. Sonja Gorjanc, Ph. D., is Assistant Professor at the Faculty of Civil Engineering (Department of Mathematics) at the University of Zagreb, the editor-in-chief of the scientific-professional journal *KoG* and the former president of the *Croatian Society for Geometry and Graphics*. Her research interests are in projective and Euclidean geometry, *Mathematica* computer graphics, curricular developments in geometry and graphics. She can be reached by e-mail: sgorjanc@grad.hr.
2. Helena Halas, M. Sc., is Research Assistant at the Faculty of Civil Engineering (Department of Mathematics) at the University of Zagreb. Her research interests are projective geometry, Euclidean and non-Euclidean planes, treated by synthetic (constructive) and analytical methods. She can be reached by e-mail: hhalas@grad.hr.
3. Ema Jurkin, Ph. D., is Assistant Professor at the Faculty of Mining, Geology and Petroleum Engineering (Department of Mathematics, Informatics and Descriptive Geometry) at the University of Zagreb, an editor of the scientific-professional journal *KoG* and the vice-president of the *Croatian Society for Geometry and Graphics* and the vice-president of *International Society for Geometry and Graphics* for the region Europe\Near East\Africa. Her research interests are projective geometry, Euclidean and non-Euclidean planes, treated by synthetic (constructive) and analytical methods. She can be reached by e-mail: ema.jurkin@rgn.hr.

INTRUIGING INTERSECTIONS AND USEFUL UNFOLDINGS

Günter BARCZIK

Erfurt University of Applied Sciences, Germany

ABSTRACT: The two geometric operations of intersecting shapes with one another and unfolding complex shapes into flat patterns have through CAD software been changed from challenge to triviality. Thus, combinations of shapes that have eluded designers for the difficulty of their handling are now at the hands of everybody who is able to use common CAD packages. We investigate what this can mean for architectural design.

Keywords: Geometry, Design, Architecture, CAD, Curriculum, Experiment.

1. A BRIEF HISTORY OF INTERSECTING SHAPES IN ARCHITECTURE

The employment of intersecting shapes in architecture has largely been confined to cuboids and cylinders, the latter already testing the limits of geometric and physical constructability. Since their conception in greek and roman architecture, groin vaults made from the intersection of two equal cylinder halves, have been the epitome of complexity in this regard (Figure 01).



Figure 01: Groin Vault: Palazzo della Ragione

the precise determination of three-dimensional intersections changed this in the 1980s. Architects like Peter Eisenman began to develop architectural designs employing numerous, sometimes excessively many intersections and overlaps - perhaps demonstrated in its most extreme form in Eisenman's never-built design for the Guardiola House in 1988 (Figures 02+03).



Figure 02: Peter Eisenman Guardiola House physical model

Only the advent of CAD software that aided

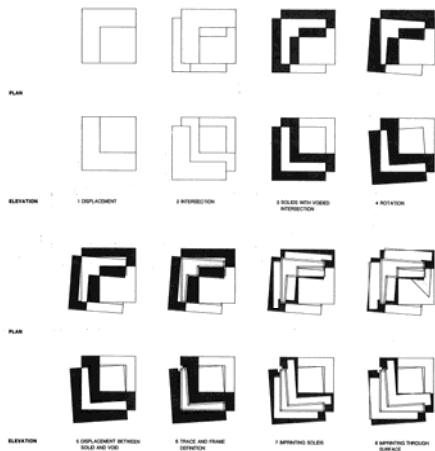


Figure 03: Peter Eisenman Guardiola House intersection studies

Where traditionally, the different sculptural and spatial segments or parts of a building were situated adjacent to one another, merely touching each other, here, in Eisenman's sculptural articulation, they merged, two elements creating an intersection as a third element - something that was shared by both and yet singular, its own self yet dependent upon its neighbours. Whether Eisenman's employment of these overlaps enriched the users' possibilities or remains a somewhat elaborate yet new form of sculptural ornament remains difficult to judge. It certainly expanded the designer's sculptural toolset and allowed him to conceive of hitherto unseen sculptural configurations. Eisenman does not intersect different forms but repeats one rectangular basic module, offsetting it in space while rotating it slightly. This subtle modification ensures that the resulting overlaps become complex but not chaotic and that the design retains a degree of discipline and sculptural soundness.

Unlike the Vitra Museum, designed and built by Frank Gehry at about the same time (in 1989). While this design is also conceived by

overlapping several distinct shapes, here the shapes are very different from one another (Figure 04). The sculptural result is much more complicated. And the actual intersections are more or less ignored by Gehry. The interest here appears to be primarily with creating a kind of spatial collage; various shapes stuck together, only creating one singular space because they overlap one another.



Figure 04: Frank Gehry Vitra Design Museum

Both Eisenman's and Gehry's projects deal more with novel outer appearances than with internal spatial possibilities.

Those are investigated in projects by the practice of Herzog de Meuron who systematically experiment with intersecting shapes and constructing buildings of various complexity and usability on the basis of these experiments. One example of this is the Jinhua Reading Space from 2007 (Figure 05): Developed from two extrusions of an irregular surface tiling intersecting one another at an angle of 90 degrees, the building is of no clear function, yet inviting playful exploration by its users and dissolving the boundaries between wall, ceiling, furniture and circulation element.



Figure 05: Herzog de Meuron Jinhua Reading Space

A second example is the Vitra House from 2011 (Figure 06).



Figure 06: Herzog de Meuron Vitra House

One simple and generic cross-section is extruded, and the resulting duct-like shape used as a module that is stacked in space in such a way that neighbouring modules intersect one another. While borrowing its degree of sculptural complexity combined with discipline and soundness from Eisenman's Guardiola House, it is nevertheless much more interested in the inside where the intersections literally open up surprising vistas and spatial relationships. Where in Gehry's Vitra Museum the intersections are ignored, here they mark the positions

of circulation and galleries.

One more example is Zaha Hadid's Zaragoza Bridge Pavillon (Figure 07) where yet another possible use of the intersections is investigated: Combining several tube-like bridge structures in such a ways that an internal spatial complexity is achieved while the structural soundness is retained.

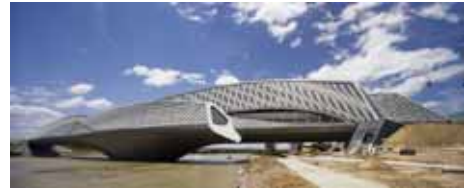


Figure 07: Zaha Hadid Zaragoza Bridge Pavillon

All five projects would have been very difficult to develop had it not been for the availability of software that actually calculated the intersections.

2. UNROLLING SURFACES

Similarly to intersecting surfaces, unrolling them is incomparably simpler with CAD software than without it. Plane cutting patterns allowing simple manufacture of surfaces can easily be produced. Again, this extends the scope of designers' possibilities much beyond what was attainable without the aid of CAD software. Unrolling surfaces and constructing complex curvatures, something that was largely ignored in architecture, had to be developed into an art form in ship-making has now become something of a commodity.

Complex unrolled surfaces become something like cutting-patterns for clothing, moving the discipline of architecture closer to that of tailoring or cloth-making, echoing Gottfried Semper's likening architecture to people's 'third skin' [1], with the natural skin people's first, and clothing people's second skin. Though while Semper keeps within his meta-

phor the fact that a skin is draped over something (natural skin as well as clothing) and a building's skin therefore draped over a structure, a surface curved or folded in an appropriate manner can itself act structurally, keeping itself upright and even supporting other elements. Therefore the employment of unrolled curved or folded surfaces opens up the possibility of freeing the understanding of architecture as a third skin from the implications of decoration that Semper's reliance on a substructure necessitated. If one regards coherence of overall structure and mutual dependability of elements as qualitative aspects of architecture, the unison of skin and structure made easier or even possible through developable surfaces surely is not unwelcome.

3. STUDENT DESIGN EXPERIMENTS

These observations above form a basis for us to conduct studies concentrating on the possible sculptural and spatial qualities intersecting surfaces can offer architectural design. Developable surfaces of increasing complexity and number are intersected, and we study the resulting spaces in terms of geometry, buildability, spatial and conceptual qualities (Figures 08-11).

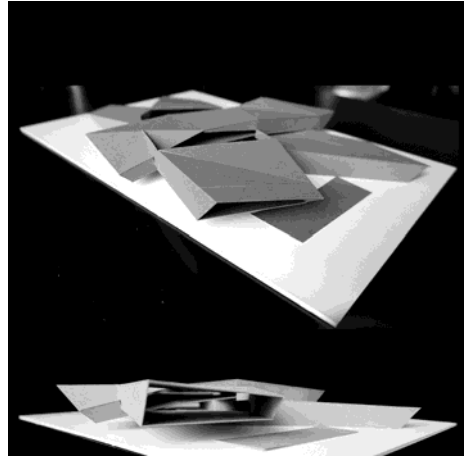


Figure 08: Student Project Andreas Beck

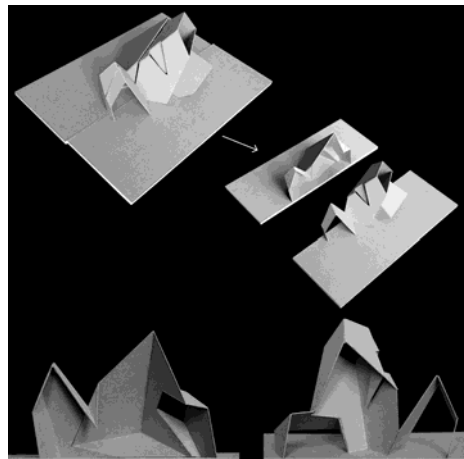


Figure 09: Student Project Andreas Beck

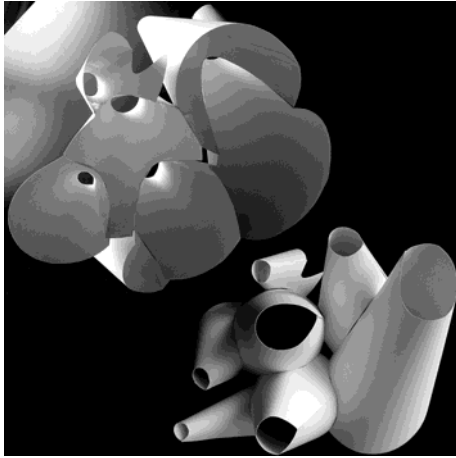


Figure 10: Student Project Lisa Dorno

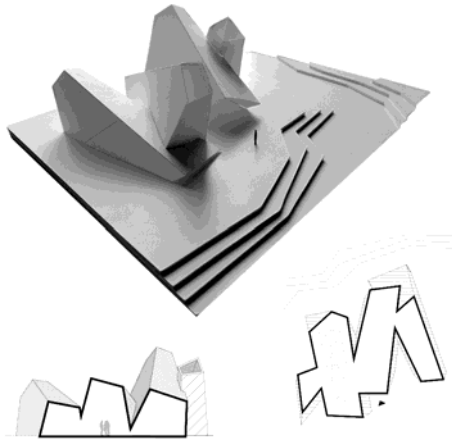


Figure 1: Student Project Madeleine Nolle

Of special interest to us are several conceptual factors: on the one hand the ambiguity of spatial regions within a shape that is the result of an intersection. When two shapes intersecting create three regions - to which shape does the new, third one, belong ? How is it connected to its neighbouring regions, being distinct yet part of a whole ? On the other hand, how are the

new, emerging regions articulated, or: which part(s) of the intersection operation are actually used, which ones discarded ? Mostly, spaces in architecture are clearly subdivided into parts. Intersections allow designers to soften transitions between regions, that is to say blur the boundaries between them.

Even with just three or four shapes intersecting, the results become very complex quickly. How can designers keep a balance between differentiated articulation and comprehensibility ?

We do not seek theoretical answers to these questions, but address them in the projects themselves by attempting to articulate various answers. We find that the new geometric possibilities open up new conceptual ones.

In terms of buildability, we develop physical from digital CAD models. Thus, we can also test the stability of the intersection results. Especially when curved parts are used, interesting possibilities emerge for constructions that use parts which are leaning against one another to achieve stability. In digital space non-developable surfaces can be applied with the same ease as developable ones. When it comes to physical production, though, this difference becomes important. Developable surfaces are much easier to produce, most importantly for us they can be build from folding patterns that the 3D modelling software can produce. We strive to not see this as a limitation but as a conceptual tool to make the designs more stringent.

All studies are conducted as part of an ongoing project with students of architecture that combines learning new geometrical skills with historical review and experimentation in terms of constructability and architectural conceptualization.

4. CONCLUSIONS AND OUTLOOK

Overall, we observe that architectural designers' spatial and sculptural repertoire can be expanded, and with it their conceptual field.

We will continue, expand and systematize our investigations with a focus on functional and structural aspects.

ACKNOWLEDGMENTS

Many thanks to all students who participated in the course in 2011-2014 as part of their Master studies in Architecture, and to the teaching assistants Max Wasserkampf and Farzad Akhavan for their help and input.

REFERENCES

- [1] G. Semper: *Der Stil in den technischen und tektonischen Künsten, oder praktische Aesthetik. Band I: Die textile Kunst für sich betrachtet und in Beziehung zur Baukunst*, Frankfurt a. M. 1860

ABOUT THE AUTHOR

Günter Barczik is Professor for Digital Design at the Architectural School of the University of Applied Sciences Erfurt and, together with Heike Matcha, co-founder and co-principle of the architectural practice HMGB Architects based in Berlin.

ITERATIVE FORM FINDING FOR THE LAYOUT OF IRREGULAR RECIPROCAL FRAMEWORK STRUCTURES

Tobias KUHLMANN¹ and Daniel LORDICK²

¹Technische Universität Berlin, Germany ²Technische Universität Dresden, Germany

ABSTRACT: This paper presents a geometrical form finding method for irregular reciprocal frame structures and its implementation as an interactive design tool. The objective for the development of the method was to apply reciprocal frames to free form designs, which are not primarily derived from the structural concept itself. We wanted a digital tool, which is easy to guide and which gives an at least close to real-time feedback. The tool is realized as an iterative algorithm in the visual programming environment Grasshopper [1]. A case study on the basis of a shading device for the terrace of a cafeteria demonstrates the applicability of the method. The algorithm and the case study are the result of the bachelor thesis of Tobias Kuhlmann and are inspired by the seminar “Bridging the Gap” by Daniel Lordick 2012 [5].

Keywords: Irregular reciprocal framework, iterative algorithm, mesh transformation, optimization, computational design, computer aided manufacturing (CAM).

1. INTRODUCTION

Reciprocal frame structures are characterized by short elements, which self-interlock cyclically to span many times the length of a single element. The usage of reciprocal frame structures in the Far East can be traced back to the 12th century as wooden roof constructions or bridges [2]. In Europe reciprocal frames are documented from the 13th century on and became rather popular through depictions in the sketchbooks of Leonardo da Vinci [8]. Information about this topic can also be found under the keywords leverworks, grillage structures, bar grids, mutually supporting structures, nex-orade, Zollinger construction or Zollbauweise and lamella structure [3]. Although reciprocal frames do not play a significant role in the building industry, there can be found nice applications in domes and also in the work of the artist Rinus Roelofs [7].

In recent times there have been some attempts to study reciprocal frames in connection with parametric modeling and computer aided manufacturing (CAM), which allow for a greater variety in structural forms and for indi-

vidually shaped bars [e.g. 6, 9]. The paper “Reciprocal Frames – Teaching Experiences” [10] in addition shows an approach to irregular reciprocal frames (Fig. 1). Irregular means that the cells, formed by the self-interlocking bars, are surrounded by varying numbers of bars overlapping in dissimilar angles.



Figure 1: Irregular reciprocal framework at the ETH Zurich [8], photograph taken from: <http://eat-a-bug.blogspot.de/2011/08>.

We suggest a new approach, where irregular reciprocal frame structures can easily be obtained and guided from a basic mesh. Thus we can handle structures above any desired floor plan. Our paper is divided into four sections.



Figure 2: Irregular reciprocal framework proposed for a canopy in front of the cafeteria at the Architecture Building, Berlin (scale model).

First, regular reciprocal frames are studied; second, generally working algorithms are developed on basis of the regular structures; third, the design options are explored and expanded and the algorithms are refined; forth, the new tool is applied to an actual design task (Fig. 2).

2. ANALYSIS OF RECIPROCAL FRAMES

We start with the analysis of regular reciprocal frames as they appear with a great variety in the work of the mathematician and artist Rinus Roelofs [7]. In his patterns there are always two kinds of cells (Fig. 3): one kind surrounded by bars, which rest clockwise on the adjacent bars and the other kind, where the bars rest counter-clockwise.



Figure 3: Pattern from Rinus Roelofs [7]; superposition with arrows to label clockwise and counter-clockwise cells.

Connecting the centers of all cells of the same kind in a way that creates polygons corresponding to the other kind of cells, results in two basic meshes. The two meshes characterize the underlying pattern of the reciprocal framework. One mesh is given in Figure 4. Both meshes can be read as limits of the reciprocal framework, if all bars of the reciprocal framework are rotated in the same direction proportionally. The pivots of the rotations are the intersections of the bars with the edges of the respective mesh. Obviously, the lengths of the bars have to be adjusted correspondingly.

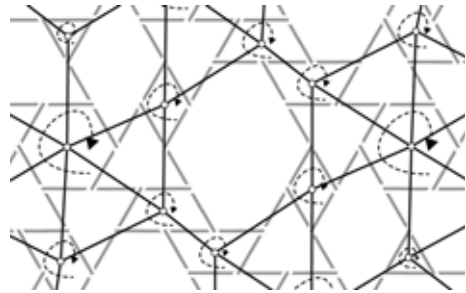


Figure 4: Basic mesh derived from the clockwise cells.

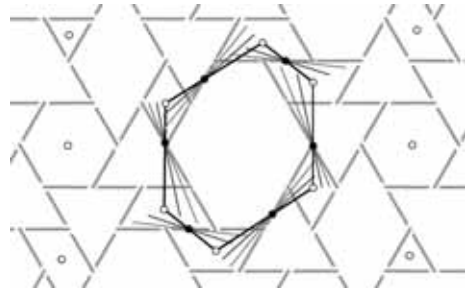


Figure 5: Transformation of a framework cell towards a polygon of the basic mesh.

To give an example, we observe the large hexagonal cell in Figure 5: All adjacent cells of the reciprocal frame, which contain vertices of the basic mesh, will collapse, which means, will be closed step by step during the proportional rotation of the bars. During their motion

the bars cover a continuous variety of positions, in which other valid reciprocal frameworks with decreasing static performance exist. In the limit the reciprocal framework will coincide with the edges of the basic mesh and thus degenerate.

3. CREATION OF IRREGULAR RECIPROCAL FRAMEWORKS

The observations of the previous chapter lead to the conjecture that the reverse motion can transform a given basic mesh into a suitable reciprocal framework. In fact the motion may lead to a range of possible solutions, depending on the chosen parameters. In principle, all solutions can be obtained with the same operational steps, which are described as follows.

With an arbitrary basic mesh given, first the algorithm has to take care of the question: which bars are adjacent to each other? During this operation also the support points at the borders of the mesh are distinguished from the inner vertices, where the missing cells of the reciprocal framework shall emerge. Now the general orientation of the cells has to be set: clockwise or counter-clockwise.

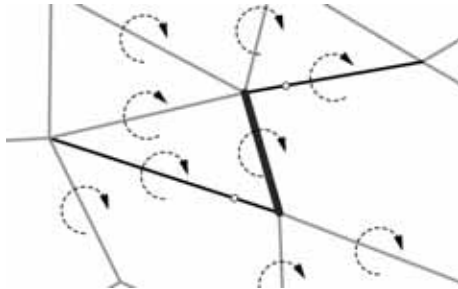


Figure 6: Adjacent lines (black) with target points for the central bold line; the local motion of each bar is set clockwise.

With this setting every endpoint of an edge gets assigned to one definite adjacent line, which will serve as a support (Fig 6). Next, there has to be determined a target point on the adjacent line, where the endpoint of the first bar

shall rest. This can be done generally by specifying a division ratio, to give a simple example.

The transformation from the basic mesh towards a reciprocal framework is computed within an iterative process, during which every endpoint of the mesh-edges moves towards a respective target point. Because of the simple fact that during the motion the adjacent edges and thus the target points are likewise moved, the transformation has to run through several iteration steps, until the endpoints reach the target points at a desired accuracy. – Please note that during the iteration steps the bars do not necessarily rotate around a certain pivot point. In this aspect the algorithm differs from the previously made analysis (Fig 7).

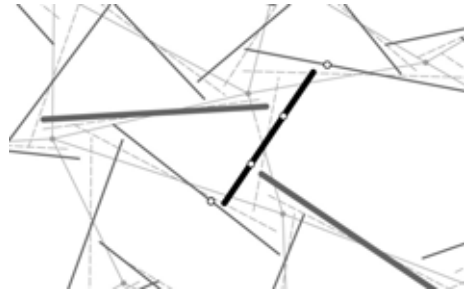


Figure 7: Motion during iteration.

During the generation of a reciprocal framework with this method, an interesting observation can be made: if the basic mesh is not planar but has support points at different heights, the algorithm will approximate an optimized surface in a sort of relaxation process. This implies, the reciprocal framework will be close to a minimal surface, which spans between the supporting points. This property facilitates the design of efficient load-bearing structures.

4. DESIGN PARAMETERS AND REFINEMENT OF THE METHOD

Until here, the main drawback of the method is that during the iteration the reciprocal framework usually deviates considerably from the

primary layout of the mesh in the floor plan (Fig. 8). This makes a formal design process with a clear guidance almost impossible. But this issue can easily be resolved by allowing the target positions to move along the adjacent bars within a certain range. We define a target zone rather than a target point. Now the algorithm needs less time to reach a reasonable solution and the frames correspond much better with the intended design (Fig. 9).

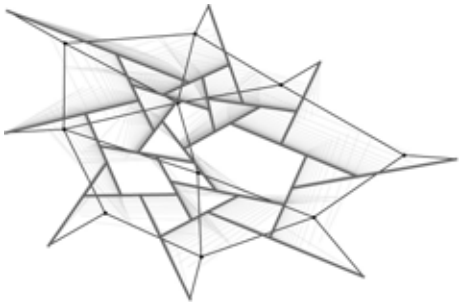


Figure 8: Strong deviation from the basic mesh caused by fixed target points.

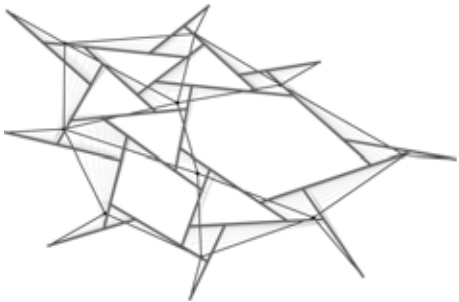


Figure 9: Good correspondence of the framework with the basic mesh by using target zones.

Another objective was to gain domelike forms with a curvature that lies beyond the effects of the previously described relaxation. This can be achieved with an additional vertical gap applied to each target position. With the definition of locally different values for all target positions, the vertical growth of the form can be steered accurately (Fig. 10).

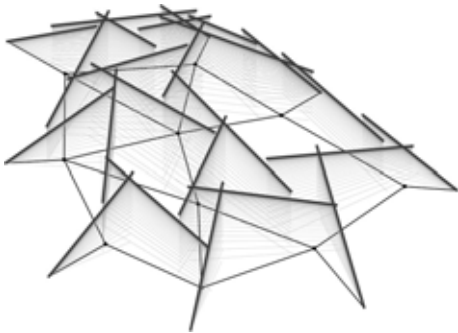


Figure 10: Creation of domelike structures.

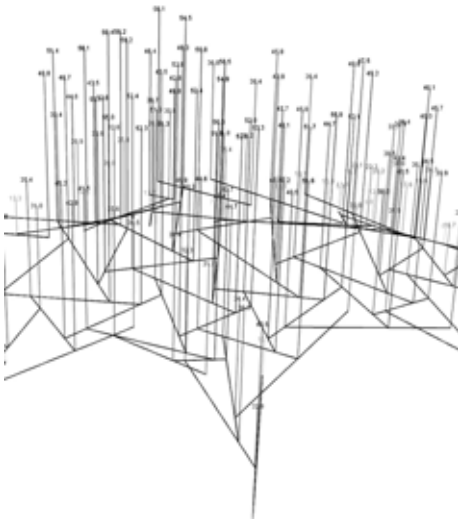


Figure 11: Static evaluation.

Up to this point, the design tool is without any method to retrieve information about the static behavior of the structure. This was upgraded by adapting a strategy given by Kohlhammer and Kotnik [4]: Planar reciprocal frameworks can be analyzed by simply considering the proportions of the distances between the supporting points on the elements. An optical feedback is presented with the help of vertical lines at all of the joints (Fig. 11). The

lengths of the lines correspond to the forces, which act in opposition to the applied dead loads in the supports. This allows for a first qualitative static evaluation in real-time. Now the resulting relative load distribution can be taken as an extra input for the optimization of the structure.

5. DESIGN OF A CANOPY

In order to test the design tool, a canopy for the cafeteria of the Institute of Architecture of the Technical University of Berlin has been developed. The cafeteria has an outdoor area, which on the one hand adjoins a deeper situated garden and on the other hand, is oriented to the forecourt of the Architecture Building at the Ernst-Reuter-Platz. Only a few spots of the existing structure are suitable for bearing the pavilion-like canopy. We finally were limited to bearing points at the edge to the garden and at the facade.

The motivation for the use of a reciprocal framework was that thus the wide span structure could be divided into relatively small elements and fabricated with the CAM facilities of the institute.

By design, the proposed reciprocal frames permit no cantilevers. The edges of the structure are concave in plan and are treated as expanded polygonal cells (Fig. 12). With this in mind the basic mesh of the reciprocal frames was defined (Fig. 13).

The basic mesh is not generated parametrically, but designed manually. It follows the idea to achieve an alternation of wide-open cells and more closed parts. This shall emphasize a casual atmosphere in front of the cafeteria. Moreover, it guarantees both, shaded areas and a good distribution of indirect light under the canopy. Although the algorithm allows for any kind of polygons, with respect to the shading devices the large cells of the framework are mainly triangles (Fig. 14, 15, 17).

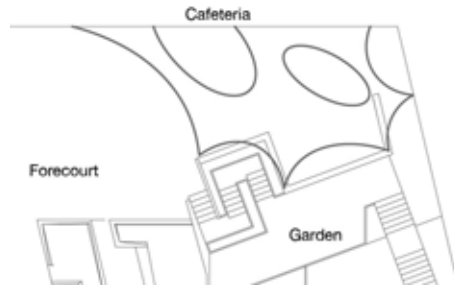


Figure 12: Area of the canopy with concave borders.

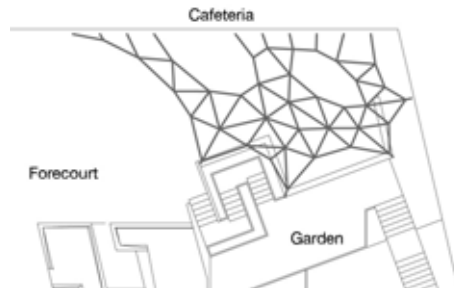


Figure 13: Basic mesh.



Figure 14: Reciprocal frames of the canopy.

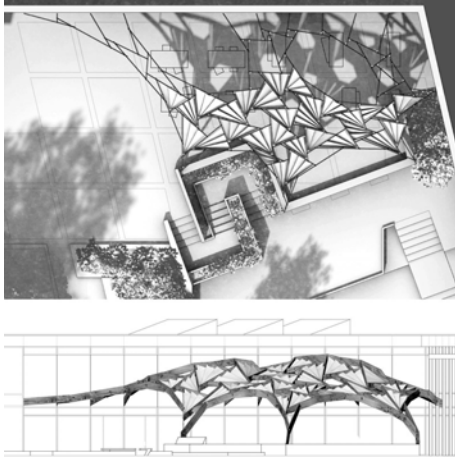


Figure 15: Plan and elevation of the canopy.

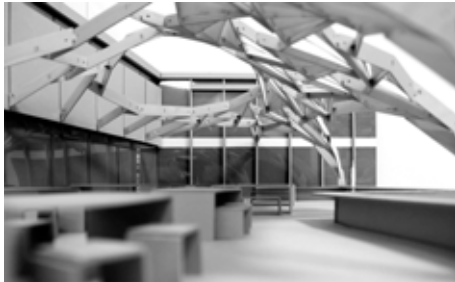


Figure 16: Impression of the scale model.



Figure 17: Detail of the scale model.

A typical key aspect of the construction was the design of a universally applicable concept for the joints, which satisfies any condition and angle between the individual bars. The challenge was to find a solution that fits at each connection without weakening the adjacent elements. The formation of the components and details had also to suit our preferred production technology: the processing of engineered wood with a 3-axis CNC milling machine.

As a proof of concept a scale model was created with a laser-cutter (Fig. 16, 17).

6. CONCLUSIONS

If acceptable control of the result is required, and this is not asked too much, the cyclical dependencies in reciprocal frameworks are challenging for any kind of optimization and transformation process.

The concept of the presented algorithm is to consider the creation of the reciprocal framework as a geometrical transformation from an initially designed two- or three-dimensional basic mesh. The basic mesh can be generated rather intuitively and can then be considered as one of two conditions, under which the desired reciprocal framework degenerates. A suitable self-interlocking structure now is obtained by the definition of the distribution of the nodes along the bars or by the definition of individual proportion values for each element.

The creation of the reciprocal framework is treated as an iterative optimization during which each endpoint of the bars attempts to reach the required destination. The algorithm allows the structure to meet different requirements like element proportions, cell typology, precise three-dimensional element intersections, or predefined vertical spacing. The iterative process makes it easy to observe the transformation process for the adjustment of design decisions and to record animations for a better understanding.

During the creation of a canopy the parametric design tool for the reciprocal framework was validated successfully. Manual work was only needed for the layout of the basic mesh.

During the layout of the mesh, the reciprocal framework adapted to all changes in close to real-time. Even the refinement of the heights of the structure succeeded with ease, without affecting the static capability.

In a further development, the algorithm could introduce irregular reciprocal frameworks to autonomous cyclic structures as proposed by the artist Rinus Roelofs.

ACKNOWLEDGMENTS

Second supervisor of the bachelor thesis of Tobias Kuhlmann was Prof. Dr.-Ing. Klaus Rückert, Chair of Design and Structures, Department of Architecture at the Technische Universität Berlin.

REFERENCES

- [1] Algorithmic Modeling for Rhino. URL: <http://www.grasshopper3d.com>. Accessed: June 2014.
- [2] V. Bertin: *Hebelstabwerke*. In: Arch+ 159/160, Zeitschrift für Architektur und Städtebau (2002): 150-153.
- [3] V. Bertin: Leverworks. URL: <http://www.vitobertin.hk/lw/index.html> Accessed: June 2014.
- [4] T. Kohlhammer; T. Kotnik: Systemic Behaviour of Plane Reciprocal Frame Structures. In: *Structural Engineering International*, 1/2011. 80-86.
- [5] D. Lordick and C. Spliid Høgsbro: Bridging the Gap. In: *Rethinking Prototyping*. Proceedings of the Design Modelling Symposium Berlin 2013. Berlin : Universität der Künste, 2013. 61-73.
- [6] M. Proll and A. Günther: Selfsupporting Framework. In: *Bauwelt*, Vol. 102, No. 1-2, 2011. 57.
- [7] Rinus Roelofs, artist's homepage. URL: <http://www.rinusroelofs.nl>. Accessed: June 2014.
- [8] A. Spiro: Reciprocal Frame Structures, 2010. URL: <http://spiro.arch.ethz.ch/de/research/reciprocal-frame.html>. Accessed: June 2014.
- [9] M. Tamke; J. Riiber; H. Jungjohann; M. Ramsgaard Thomsen: Lamella Flock. In: *Advances in Architectural Geometry 2010*. Wien, New York : SpringerWienNewYork, 2010. 37-48.
- [10] U. Thönnissen; N. Werenfels: Reciprocal Frames – Teaching Experiences. In: *International Journal of Space Structures* Vol. 26 No. 4, 2011. 369-371.

ABOUT THE AUTHORS

1. Tobias Kuhlmann studied architecture at the TU Berlin. He works as a freelancer for Househam Henderson Architects among others.
2. Daniel Lordick studied architecture at the TU Berlin and the Carleton University in Ottawa. He is professor for Geometric Modeling and Visualization at the TU Dresden, Institute of Geometry.

KINEMATIC EXPLORATION OF 1-DOF ORIGAMI MECHANISMS

Yves Klett and Peter Middendorf
University of Stuttgart, Germany

ABSTRACT: This paper will present a study on the kinematic behavior of non-trivial folded structures with one kinematic degree of freedom. Motivated by the need for new lightweight sandwich core structures for aircraft components, several origami tessellations have been identified as potential templates for high-performance cellular materials, so-called foldcores. To design and manufacture these tessellations, it was necessary to develop geometric models and methods to simulate their kinematic behavior. These tools allow the exploration of origami structures as idealized mechanisms. We will introduce the topic of 1-DOF rigid origami and present selected examples of interesting origami mechanisms together with the visualization of the folding process.

Keywords: Geometry, Rigid Origami, Kinematics, Engineering, Folding, Mechanism

1. INTRODUCTION

Folding a piece of paper, is a trivially simple process – at first. By combining more than one crease in an interconnected crease pattern, the folding process and the resulting structure can become arbitrarily complex. On the other hand, the concept of origami is extremely amenable to geometric analysis. Several well-know origami rules or axioms translate directly into geometrical concepts and idealizations. Two important postulates for the following experiments are that paper is assumed to be thin, and that paper does not stretch. This translates into an approximation of paper by surfaces of zero thickness which maintain isometry during all folding stages.

Folding does not only appeal to origamists and mathematicians (often in personal union) though, but has direct technological implications. Folded structures perform essential every-day life tasks from packaging to life-saving. [1]

2. ORIGAMI MEETS SANDWICH

One relatively new application for clever folding has been identified in the field of lightweight engineering. Here, so called sandwich structures allow the construction of light high-performance structures, which consist of two solid face sheets

and an inner, very light core (Figure 1). For the core part several origami tessellations have been found to offer good performance in combination with other properties and functionalities that are not provided by other state-of-the-art materials like honeycombs, foams or woods. [2]

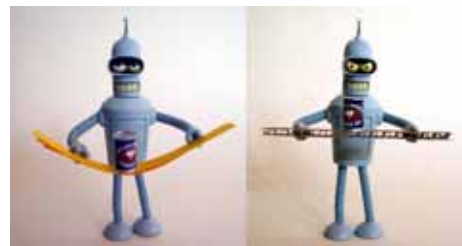


Figure 1: The sandwich principle: Clever use of the same amount of material results in superior performance.

3. RIGID ORIGAMI AS MECHANISM

To precisely design and manufacture such cellular structures (also termed foldcores, see Figure 2), it was necessary to develop geometrical models and methods to simulate the folding process. Several approaches to this have been described,

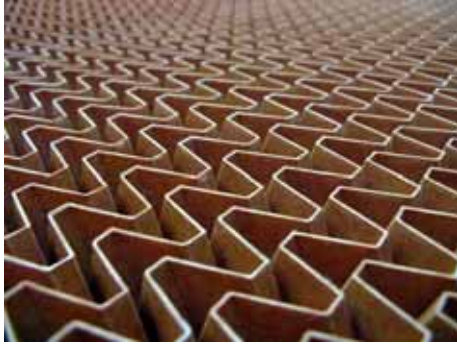


Figure 2: Foldcore sandwich hardware manufactured from Kevlar[®]-based paper.

ranging from purely analytical to finite-element based methods [3–6]. For the following analysis a geometry-based algorithm was used that allows for real-time simulation of the folding behavior of quad-based tessellations with one kinematic degree of freedom [7]. The algorithm takes a 3D tessellation as input and implicitly computes the resulting geometry for given values of the folding parameter μ (Figure 3). All faces of the given tessellation are treated as rigid, and all creases as hinges with one rotational degree of freedom.

For the presented non-trivial structures, these multiple degrees of freedom are constrained by the underlying morphology and effectively result in a mechanism with one kinematic degree of freedom, i.e. the change of one dihedral angle between two adjacent faces will affect all other dihedral angles in turn.

4. KINEMATICS OF 1-DOF RIGID ORIGAMI

Intimate knowledge of the the kinematic behavior of a given 1-DOF mechanism is essential for the successful design of both the structure itself and the manufacturing tools. Close adherence to the isometric principle during folding is especially important for materials with high stiffness and low breaking strain.

But also the kinematics of rigid origami struc-

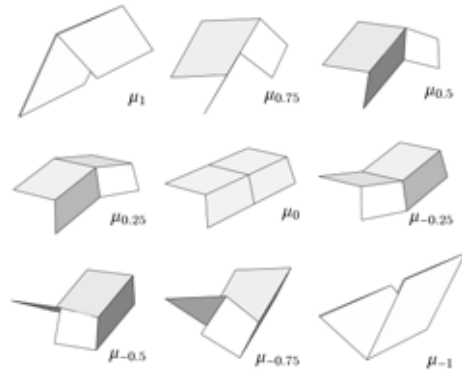


Figure 3: Folding kinematics for a single miura cell. The parameter μ describes the folding state.

tures themselves can be of great interest: Whether the resulting mechanisms are folded from sheet material or are realized by thick platelets and actual hinges [8], the same basic principles still apply for their preliminary design and very often translate directly into reality.

The following kinematic studies are all based on idealized models, ignoring physical properties and effects like elasticity, mass, gravity or friction, but have been checked for self-intersection and thus could be principally realized as functional mechanical models.

4.1 Planar movement

Planar movements are defined by coplanar movement of one or more vertices of a given mechanism. Such kinematics can be useful for technical applications, e.g. to generate straight-line mechanisms that translate circular motion to a linear one and vice versa by means of a three-dimensional structure. For planar kinematics, the coordinates of any vertex can be formulated in a compact closed form that is only depending on the flat folding pattern, the initial folding direction and the folding parameter μ [2].

For subsequent visualization the trace of selected vertices during folding will be shown in different configurations. In addition to the

crease pattern and the initial folding direction, the choice of origin and face alignment strongly influence the resulting global movement. The folding range from μ_1 to μ_{-1} was chosen to show the always present symmetry between negative and positive folding states. In a mechanical realization the switch from negative to positive μ is the most critical point, because all crease need to reverse their valley/mountain orientation at μ_0 .

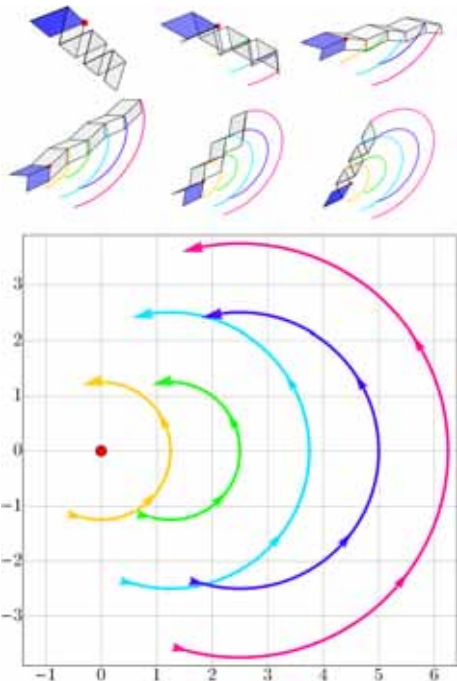


Figure 4: Planar trace of a miura cell stripe.

Figure 4 shows a strip of miura type cells which are aligned vertically parallel to the mean of the vertex normals of the blue polygons. The origin is indicated by a red dot. The resulting traces consist of pairwise offset circular segments with growing radii. An analytic solution for the resulting traces is detailed in [2].

In Figure 5 the same strip is traced with a different alignment, resulting in ellipsoid segments

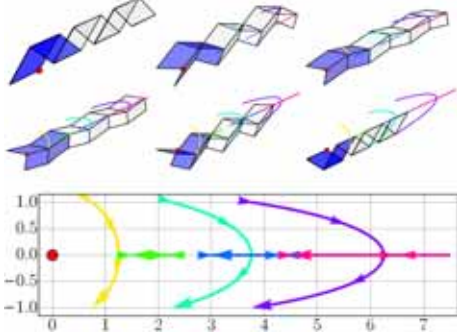


Figure 5: Trace for the strip from Figure 4 with different alignment.

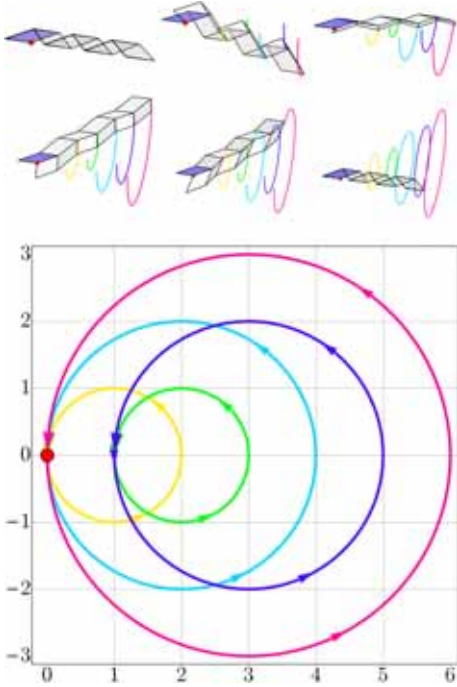


Figure 6: Same strip, different alignment.

with growing aspect ratio that are interspersed with straight linear segments with likewise grow-

ing amplitude.

Yet another alignment for the same strip is shown in Figure 6, where the vertices trace complete circles with pairwise growing radii in parallel planes. A strip with a rather arbitrarily chosen crease pattern is shown in Figure 7. The traces exhibit several crossing points. As a final example for planar curves traced by origami mechanisms, Figure 8 shows the movement of a spiralling curved strip.

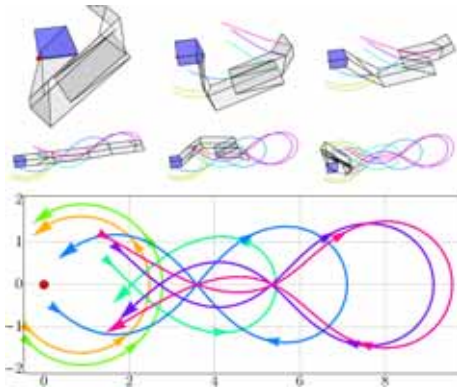


Figure 7: A random strip with multiple trace crossings.

4.2 Threedimensional movement

A much wider variety of movements becomes available for the non-planar, threedimensional case. Figure 9 shows an overlay of twenty folding states from μ_0 to μ_1 for a symmetric miura cell. In this case as well for the geometries shown in Figures 10-12 available symmetries allow a compact closed-form description of the resulting kinematics, while for geometries with torsional behavior like e.g. shown in Figure 14 the derivation becomes much more complex. This can be problematic in the context of reverse-engineered mechanisms that need to be tailored to specific movements, but does not affect the forward kinematic simulation of given geometries.

With a suitable choice of reference origin and alignment, the kinematic traces can be used to

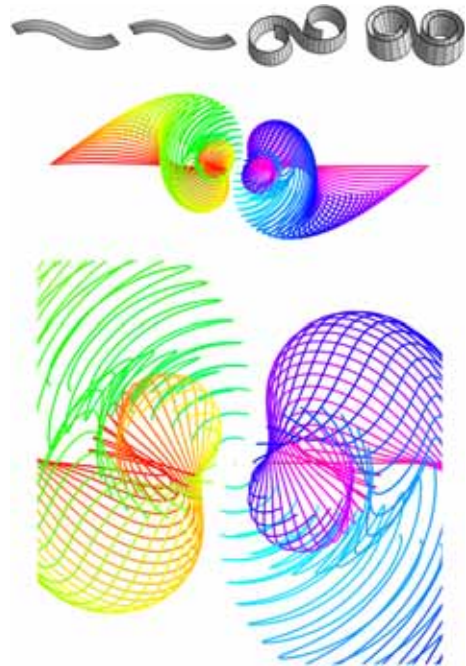


Figure 8: Planar traces of a spiralling band.

compute the folding envelope of a tessellation, which is e.g. relevant for the design of production tools. This is shown in Figure 10, where the traces are plotted with reference to the center of gravity of the structure. The same alignment is chosen for the curved, sinusoidal cell array from Figure 12, where the chosen discretized vertices trace a distinct wavy folding envelope.

For different, asymmetric alignments the three-dimensional traces can become quite complex. In Figure 13 a strip of "waterbomb" unit cells is aligned with respect to a single polygon which together with the inherent retrograde behavior of this cell type results in a an envelope with characteristically sharp reversal points.

Finally, Figure 14 shows what happens during the folding of a strip-based toroid, which exhibits a lot of torsional movement. Quite surprisingly,

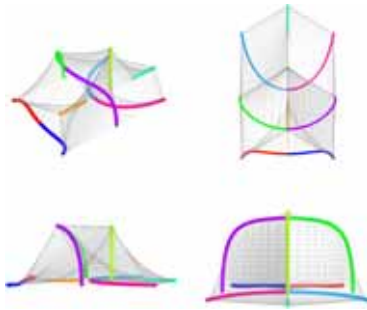


Figure 9: Overlay of 20 folding states.

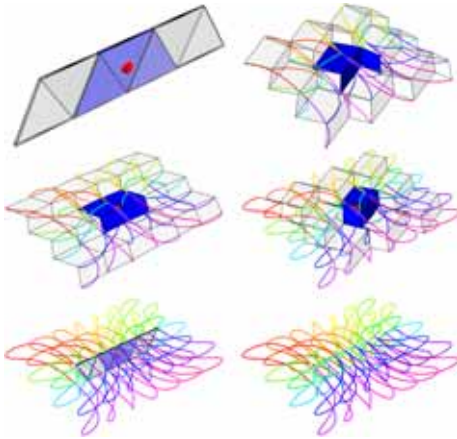


Figure 10: Vertex traces for a 3×3 miura array for from μ_1 to μ_{-1} .

this structure is still rigidly foldable without self-intersection.

In general, the perceived complexity of a set of vertex traces for a given rigid origami structure will correspond quite nicely with its manufacturing effort. Simple, clear traces as in Figure 10 indicate the potential for an automated production process of industrially relevant structures, while complex structures as shown in Figure 14 will most likely remain in the domain of skilled folders for the near future.

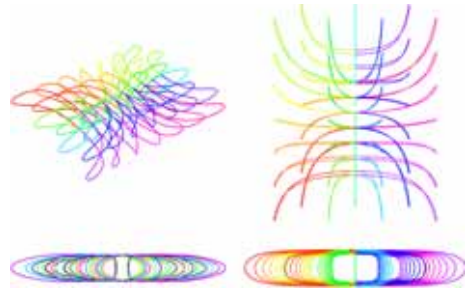


Figure 11: The traces from Figure 10 shown from different angles.

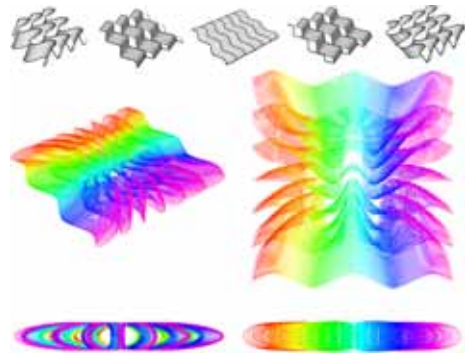


Figure 12: Top row: Folding of a sinusoidal curved structure. Bottom: Kinematic envelope as traced by discretized vertices.

5. CONCLUSIONS

We have demonstrated the kinematic behavior of several origami tessellations of differing complexity. The visualization of vertex traces during the folding process using an implicit folding algorithm reveals the complex and sometimes surprising behavior of a small number of selected structures. Both planar and three-dimensional examples were presented that can be used to generate well-defined and predictable movements using just one actuator.

The presented method has been put to good use in the design of high-performance lightweight structures, but is by no means restricted to this ap-

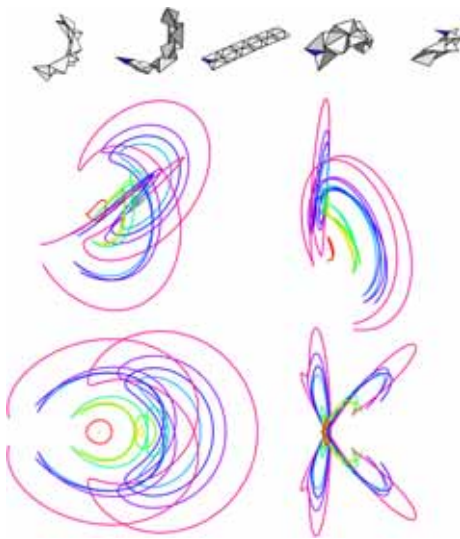


Figure 13: Midpoint traces produced by folding a strip of "waterbomb" cells.

plication. In the field of tailored mechanisms and in education, computational and applied origami can prove to be a very interesting and entertaining venture with a low entry barrier – the next piece of paper is *always* around.

REFERENCES

- [1] Y. Klett and K. Drechsler. Designing technical tessellations. In P. Wang-Iverson, R. J. Lang, and M. Yim, editors, *Origami⁵*, number ISBN 1568817142, pages 305–322. CRC Press, 2011.
- [2] Yves Klett. *Auslegung multifunktionaler isometrischer Faltstrukturen für den technischen Einsatz*. ISBN 978-3-8439-1025-5, Universität Stuttgart, 2013.
- [3] Yves Klett and Klaus Drechsler. Design of multifunctional folded core structures for aerospace sandwich applications. In *1st CEAS European Air and Space Conference*, 2007.

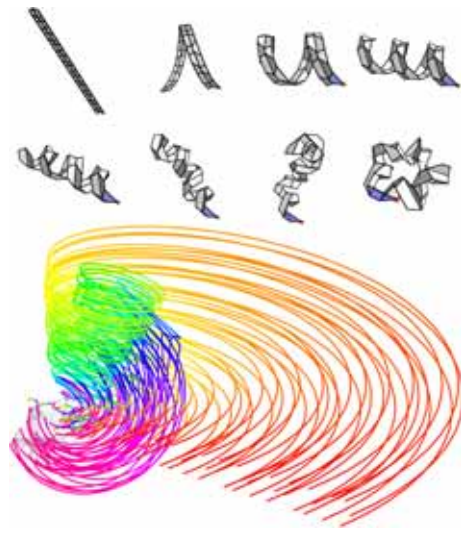


Figure 14: Folding traces of a strip-based toroid from μ_0 to μ_1 . Colors are assigned per μ , so all points of one color represent one folding state.

- [4] Tomohiro Tachi. Simulation of rigid origami. In *Origami⁴ 4th International Conference on Origami in Science, Mathematics and Education*, pages 175–187, 2009.
- [5] Marcelo A. Dias, Levi H. Dudte, L. Mahadevan, and D. Satnagelo. Geometric mechanics of curved crease origami. *arXiv.org*, 2012. URL <http://arxiv.org/pdf/1206.0461v2.pdf>.
- [6] M. Schenk and S. D. Guest. Origami folding - a structural engineering approach. In P. Wang-Iverson, R. J. Lang, and M. Yim, editors, *Origami⁵*, pages 291–303. CRC Press, 2011.
- [7] Y. Klett. Realtime rigid folding algorithm for quadrilateral-based 1-DOF tessellations. In *ASME International Design Engineering Technical Conferences, DETC2013-12659*, 2013.

- [8] Tomohiro Tachi. Rigid-foldable thick origami. In P. Wang-Iverson, R. J. Lang, and M. Yim, editors, *Origami*⁵, number ISBN 1568817142, pages 253–263, 2011.

ABOUT THE AUTHORS

1. Yves Klett leads the sandwich group at the Institute of Aircraft Design in Stuttgart (klett@ifb.uni-stuttgart.de).
2. Peter Middendorf is the director of the Institute of Aircraft Design, University of Stuttgart (middendorf@ifb.uni-stuttgart.de).

KINEMATICS OF AN OVERCONSTRAINED MOBILE RRRCR LOOP

Jesús CERVANTES-SÁNCHEZ¹ and Paul J. ZSOMBOR-MURRAY²

¹Universidad de Guanajuato, México

²McGill University, Canada

ABSTRACT: Displacement, velocity and acceleration analysis of a novel 5 link, 5 joint closed chain mechanism is carried out. After introducing the related families of 4, 5 and 6R arrangements it is shown how formulation in an ideally chosen frame presented in 3 descriptive geometric views provides for simple, systematic analysis. The essential mechanism model is a planar “slider-crank” where the “wrist-pin” moves in pure translation such that all points on this rigid body trace parallel, coaxial elliptical paths; sections of coaxial right cylinders. The wrist pin “slides” on a generator of one of these. In terms of the crank angle displacement analysis yields a quartic solution. In terms of an input angle measured around the cylinders, the solution is quadratic. Acceleration analysis is facilitated by using the curvature centre curve of the elliptical cylinder section on which the wrist-pin moves.

Keywords: Mechanism, descriptive geometry, overconstraint, kinematics, analysis

1. INTRODUCTION

The mechanism shown in Fig. 1 is a special case of a family of mobile 6R (revolute) simple closed chains that has been the subject of extensive study. Although there are more recent references a monumental work [3], by Peter Dietmaier of TU-Graz, devotes over 50 pages to movable 6R loops. Therein is cited the earliest example invented by Sarrus’ [7] and another extensive review article on the topic by Baker [1]. In this regard Dietmaier also mentions contributions by Mavroidis and Roth [5, 6], Waldron [8] and three of Wohlhart’s [9–11]. These previous investigations, except for [5, 6] that include topologies very similar to that addressed in this paper, are largely devoted to 6R loops that are derived by combining or augmenting simpler over-constrained R-jointed loops, *i.e.*, the 4R Bennett’s mechanism [2], of which there is but one type, and Goldberg’s [4] or other 5R mechanisms, which have been studied more thoroughly than have 6R loops. In the example at hand, composed of rigid bodies labeled 1, ..., 5 in Fig. 1, the cylindrical or C-joint may be regarded as comprising a degenerate sixth link of zero length where the prismatic or P-joint element of the C-joint has its axis on

an absolute line. Another special feature is provided by a serial RRR subchain whose axes AOQ , BI and CD are mutually parallel. It is easy to see that with a general skew axis RRR triad the RRRCR loop would be immobile, quite apart from a Chebyshev-Grubler-Kützbach mobility computed as zero, because the triad is sufficient to place a specific *point* (a 3-parameter object) but *not* a line (a 4-parameter object). A brief geometric approach of its position analysis will be outlined first with angular input angle α applied to R-axis AOQ , called direct kinematics or DK, then θ to R-axis FG , called inverse kinematics or IK. Before proceeding one notes that during motion the intermediate link CDE undergoes a pure translation because directions of lines CD and DE remain fixed. All points on rigid body CDE describe identical elliptical paths. The one on point D traces the section of the right cylinder, swept by generator on DE as it revolves about axis FG at radius $r_F = EF$, cut by plane $z = 0$, selected as such when setting up an ideal frame to simplify otherwise unwieldy mathematical expressions.

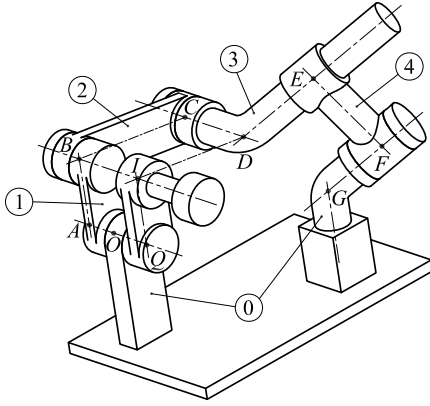


Figure 1: A 1dof Overconstrained Mechanism

2. GEOMETRY

In order to provide the simplest possible input/output motion model of this mechanism one may adopt the three view configuration shown in Fig. 2 and, in the interest of clarity, indulge in a bit of repetition. The principal plane $z = 0$ is taken to lie on points Q, I, D . The partial mechanism projected thereon is a special planar 4-bar with crank link QI , length r_Q , coupler ID , length r_I , such that the “wrist-pin” D is constrained to move on an ellipse. Note the crank-angle α (input) and the coupler angle β . The elliptical path is described by the section of a right cylinder, radius r_F , inclined to plane $z = 0$ at ϕ –the device shown in Fig. 1 tipped so AOQ , BI and CD all point upward–, and plane $z = 0$ itself. Notice that point H is the intersection of the cylinder axis, on points FG , and the plane while the point D , also on $z = 0$, is on a cylinder generator DE . The output angle θ is measured from a ray on the cylinder axis and normal to $y = 0$. Without loss in generality, placement and orientation of the 4-bar and the cylinder are such that the crank pin circle k_Q , the coupler circle k_I –where $I(x_I, y_I) = (r_Q \cos \alpha, r_Q \sin \alpha)$ – and the el-

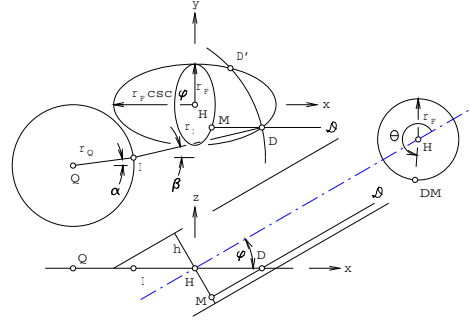


Figure 2: Geometric Input/Output Model

lipse equations k_H are written as Eq. 1.

$$\begin{aligned} k_Q : (x - x_Q)^2 + (y - y_Q)^2 - r_Q^2 &= 0 \\ k_I : (x - x_I)^2 + (y - y_I)^2 - r_I^2 &= 0 \\ k_H : Bx^2 + Ay^2 - AB &= 0 \end{aligned} \quad (1)$$

The ellipse is origin centred on H with its major axis along the principal x -axis. $A = a^2$ where a is length of major axis. Similarly $B = b^2$ where b is the minor axis length. Point D can be expressed as Eq. 3 in terms of the output angle θ , the cylinder slope angle ϕ and the plane $z = 0 \equiv z_O$. First, Eq. 2, the radial Plücker coordinates of the generator \mathcal{D} on point D are derived according to [12]. Then point D is located as $D = \mathcal{D} \cap z_O$.

$$\begin{aligned} \begin{vmatrix} 1 & -r_F \sin \theta \sin \phi & r_F \cos \theta & r_F \sin \theta \\ 0 & \cos \phi & 0 & \sin \phi \end{vmatrix} \rightarrow \\ \mathcal{D} \{d_{01} : d_{02} : d_{03} : d_{23} : d_{31} : d_{12}\} \\ \mathcal{D} \{\cos \phi : 0 : \sin \phi : r_F \cos \theta \sin \phi \\ : r_F \sin \theta : -r_F \cos \theta \cos \phi\} \end{aligned} \quad (2)$$

$$\begin{aligned} D = \mathcal{D} \cap z_O, \quad d_i = \sum_{j=0}^3 d_{ij} Z_{Oj} \\ z_O \{Z_{O0} : Z_{O1} : Z_{O2} : Z_{O3}\} = \{0 : 0 : 0 : 1\} \end{aligned}$$

$$\begin{aligned} d_0 &= +d_{01}Z_{O1} + d_{02}Z_{O2} + d_{03}Z_{O3} \\ d_1 &= -d_{01}Z_{O0} + d_{12}Z_{O2} - d_{31}Z_{O3} \\ d_2 &= -d_{02}Z_{O0} - d_{12}Z_{O1} + d_{23}Z_{O3} \\ d_3 &= -d_{03}Z_{O0} + d_{31}Z_{O1} - d_{23}Z_{O2} \end{aligned}$$

$$\begin{aligned}
& D\{d_0 : d_1 : d_2 : d_3\} \\
& = \{\sin \phi : -r_F \sin \theta : r_F \cos \theta \sin \phi : 0\} \\
& \equiv D(x_D, y_D) \\
& = (-r_F \sin \theta \csc \phi, r_F \cos \theta)
\end{aligned} \tag{3}$$

3. POSITION ANALYSIS

Direct position analysis presumes a given value of crank angle α . This locates the crank pin at point I .

$$x_I = r_Q \cos \alpha, \quad y_I = r_Q \sin \alpha$$

A quartic resultant in x , Eq. 4, is obtained by eliminating y between circle k_I and ellipse k_H , given in Eq. 1.

$$\begin{aligned}
& c_4 x^4 + c_3 x^3 + c_2 x^2 + c_1 x + c_0 = 0 \\
& c_4 = (A - B)^2, \quad c_3 = -4A(A - B)x_I \\
& c_2 = 2A[(3x_I^2 + y_I^2 + B - r_I^2)A \\
& \quad - (x_I^2 - y_I^2 + B - r_I^2)B] \\
& c_1 = -4A(x_I^2 + y_I^2 + B - r_I^2) \\
& c_0 = A^2[(x_I^2 + 2y_I^2 - 2r_I^2)x_I^2 \\
& \quad + (y_I^2 - 2r_I^2)y_I^2 \\
& \quad + (2x_I^2 - 2y_I^2 + B - 2r_I^2)B + r_I^4]
\end{aligned} \tag{4}$$

where (x, y) implies $D(x_D, y_D)$. Some mechanism architectures might admit four real intersections between k_I and k_H . The one shown in Fig. 2 admits two, labeled D and D' . To match four values of y to x without ambiguity one may use Eq. 5 which is linear in y . It is obtained by eliminating y^2 between k_I and k_H in Eq. 1.

$$\begin{aligned}
& 2Ay_I y - A[(x - x_I)^2 + y_I^2 - r_I^2] \\
& + Bx^2 - AB = 0
\end{aligned} \tag{5}$$

Output angle, thus completing direct position analysis, is unambiguously obtained with Eq. 3. In inverse position analysis one is given input angle θ which immediately locates $D(x_D, y_D)$ again via Eq. 3. The intersection of circle k_Q from Eq. 1 and circle $k_D : (x - x_D)^2 + (y - y_D)^2 - r_I^2 = 0$

yields a quadratic univariate in x produced by eliminating y from k_Q with the difference $k_D - k_Q$ which is linear in both x and y . Eq. 6 sums this up.

$$\begin{aligned}
& k_D - k_Q : \\
& 2(x_Q - x_D)x + 2(y_Q - y_D)y \\
& + x_D^2 - x_Q^2 + y_D^2 - y_Q^2 \\
& - r_I^2 + r_Q^2 = 0 \\
& \rightarrow Px + Qy - R = 0 \\
& (P^2 + Q^2)x^2 + 2(PQy_Q - PR - Q^2x_Q)x \\
& - Q^2r_Q^2 + R^2 + Q^2x_Q^2 \\
& + Q^2y_Q^2 - 2RQy_Q = 0
\end{aligned} \tag{6}$$

4. VELOCITY ANALYSIS

Direct analysis presumes that results of position analysis are available as is angular velocity $\dot{\alpha}$. With definitions $x_{ID} = x_D - x_I$ and $y_{ID} = y_D - y_I$ one may immediately express the following velocities

$$\begin{aligned}
\mathbf{v}_I &= \begin{bmatrix} -\dot{\alpha}y_I \\ \dot{\alpha}x_I \\ 0 \end{bmatrix}, \quad \mathbf{v}_D = \mathbf{v}_I + \mathbf{v}_{D/I} \\
\mathbf{v}_{D/I} &= \begin{bmatrix} -\dot{\beta}y_{ID} \\ \dot{\beta}x_{ID} \\ 0 \end{bmatrix} \\
\mathbf{v}_D &= k_1 \begin{bmatrix} -Ay_D \\ Bx_D \\ 0 \end{bmatrix} = \begin{bmatrix} \dot{x}_D \\ \dot{y}_D \\ \dot{z}_D \end{bmatrix}
\end{aligned} \tag{7}$$

where $\dot{\beta}$ is the angular velocity of the coupler, k_1 is a constant of proportionality and the vector following it in the direction of the ellipse tangent at D . This tangent can be conveniently obtained as the polar line g_P at D .

$$\begin{aligned}
g_P : \begin{bmatrix} 1 & x & y \end{bmatrix} \begin{bmatrix} -AB & 0 & 0 \\ 0 & B & 0 \\ 0 & 0 & A \end{bmatrix} \begin{bmatrix} 1 \\ x_D \\ y_D \end{bmatrix} \\
= -AB + Bx_D x + Ay_D y = 0
\end{aligned}$$

Notice that coefficients $\{Bx_D : Ay_D\}$ are direction numbers *normal* to the tangent line. Solution

of this velocity system in detached coefficient homogeneous form is written as Eq. 8.

$$\begin{bmatrix} -\dot{\alpha}y_I & -y_{ID} & -Ay_D \\ \dot{\alpha}x_I & x_{ID} & Bx_D \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \beta_1 \\ k'_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (8)$$

$$\begin{aligned} \lambda_1 &= Ax_{ID}y_D - By_{ID}x_D \\ \beta_1 &= \dot{\alpha}(By_Ix_D - Ax_Iy_D) \\ k'_1 &= \dot{\alpha}(x_Iy_{ID} - y_Ix_{ID}) \\ \beta &= \beta_1/\lambda_1 \\ k_1 &= k'_1/\lambda_1 \end{aligned}$$

Thus the angular velocity $\dot{\beta}$ of the coupler and the “wrist pin” velocity of D is determined. Since the plane h of the circular trajectory of E can be moved onto H due to pure translation of the intermediate link, to get θ the angular velocity of the output link one may write Eq. 9 which has a built in computational check. Any pair of the three equations may be used to determine the two unknown parameters $\dot{\theta}, k_2$.

$$\begin{aligned} \mathbf{v}_D &= \mathbf{v}_E + \mathbf{v}_{D/E} \\ \mathbf{v}_E &= \dot{\theta} \begin{bmatrix} \cos \phi \\ 0 \\ \sin \phi \end{bmatrix} \times \begin{bmatrix} x_M \\ y_M \\ z_M \end{bmatrix} \\ &= \begin{bmatrix} -\dot{\theta} \sin \phi y_M \\ \dot{\theta} (\sin \phi x_M - \cos \phi z_M) \\ \dot{\theta} \cos \phi y_M \end{bmatrix} = \begin{bmatrix} \dot{x}_E \\ \dot{y}_E \\ \dot{z}_E \end{bmatrix} \quad (9) \\ \mathbf{v}_{D/E} &= k_2 \begin{bmatrix} \cos \phi \\ 0 \\ \sin \phi \end{bmatrix} \end{aligned}$$

where $M(x_M, y_M, z_M)$ is given by a line/plane intersection.

$$\begin{aligned} M &= \mathcal{D} \cap h, \quad j_i = \sum_{j=0}^3 d_{ij} H_j \\ h \{H_0 : H_1 : H_2 : H_3\} \\ &= \{0 : \cos \phi : 0 : \sin \phi\} \end{aligned}$$

$$\begin{aligned} m_0 &= +d_{01}H_0 + d_{02}H_2 + d_{03}H_3 \\ m_1 &= -d_{01}H_0 + d_{12}H_2 - d_{31}H_3 \\ m_2 &= -d_{02}H_0 - d_{12}H_1 + d_{23}H_3 \\ m_3 &= -d_{03}H_0 + d_{31}H_1 - d_{23}H_2 \end{aligned}$$

$$\begin{aligned} &= 1 \\ &= -r_F \sin \theta \sin \phi = x_M \\ &= r_F \cos \theta = y_M \\ &= r_F \sin \theta \cos \phi = z_M \end{aligned}$$

To sum up the unknowns in Eq. 9 are $\dot{\theta}, k_2$ going forward and k_1, k_2 when $\dot{\theta}$ is input. In the latter case \mathbf{v}_D is obtained with Eq. 7 and getting $\dot{\beta}, \dot{\alpha}$ from there is straightforward.

5. ACCELERATION ANALYSIS

In this section angular accelerations $\ddot{\alpha}$ and $\ddot{\theta}$ will be imposed, as direct and inverse inputs, together with position and velocity data, computed previously with position and velocity analysis procedures, respectively, to obtain \mathbf{a}_D . Components $\mathbf{a}_{nD} + \mathbf{a}_{tD} = \mathbf{a}_D$, in normal and tangential directions at point D relative to the ellipse, are used and the computation is summed up in Eq. 10.

$$\begin{aligned} \mathbf{a}_D &= \mathbf{a}_I + \mathbf{a}_{D/I}, \quad \mathbf{a}_I = \mathbf{a}_{nI} + \mathbf{a}_{tI} \\ \mathbf{a}_{D/I} &= \mathbf{a}_{nD} + \mathbf{a}_{tD} \mathbf{a}_M = \mathbf{a}_E = \mathbf{a}_{nM} + \mathbf{a}_{tM} \quad (10) \\ \mathbf{a}_D &= \mathbf{a}_M + \mathbf{a}_{D/M} \end{aligned}$$

such that

$$\begin{aligned} \mathbf{a}_I &= -\frac{\mathbf{v}_I^2}{r_Q} \mathbf{r}_Q + \ddot{\alpha}_Q \times \mathbf{r}_Q \\ \mathbf{a}_{nD} &= \frac{\mathbf{v}_D^2}{\rho_D} \mathbf{n}_D, \quad \mathbf{a}_{tD} = k_3 \begin{bmatrix} Ay_D \\ -Bx_D \end{bmatrix} \\ \mathbf{a}_{nM} &= -\ddot{\theta}^2 \begin{bmatrix} x_M - x_H \\ y_M - y_H \\ z_M \end{bmatrix} \\ \mathbf{a}_{tM} &= \ddot{\theta} \begin{bmatrix} \cos \phi \\ 0 \\ \sin \phi \end{bmatrix} \times \begin{bmatrix} x_M - x_H \\ y_M - y_H \\ z_M \end{bmatrix} \\ \mathbf{a}_{D/M} &= k_4 \begin{bmatrix} \cos \phi \\ 0 \\ \sin \phi \end{bmatrix} \end{aligned}$$

and

$$\begin{aligned} \mathbf{n}_D &= -\frac{1}{\sqrt{(Bx_D)^2 + (Ay_D)^2}} \begin{bmatrix} Bx_D \\ Ay_D \end{bmatrix} \\ \rho_D &= \sqrt{\frac{(A^2 - Ax_D^2 + Bx_D^2)^3}{A^4 B}}. \end{aligned}$$

Radius of curvature ρ_D is the reciprocal of curvature κ_D at point D on the ellipse. Eq. 10 and the relations below it are sufficient to compute either forward or inverse accelerations.

5.1 Ellipse Curvature

The explicit curvature function of a standard form ellipse, parameterized on x as $\kappa(x)$, or rather its square $\kappa^2(x)$, a much simpler expression, is derived as Eq. 11.

$$\begin{aligned} Bx^2 + Ay^2 - AB &= 0 \rightarrow \\ y &= \sqrt{\frac{AB - Bx^2}{A}} \\ y'^2 &= (B^2x^2) / [A(AB - Bx^2)] \\ y''^2 &= AB / (A - x^2)^3 \rightarrow \\ \kappa^2 &= y''^2 / (1 + y'^2)^3 \\ &= A^4B / (A^2 - Ax^2 + Bx^2)^3 \end{aligned} \quad (11)$$

It is noted with satisfaction that radius of curvature $\rho = 1/\kappa$, a key parameter in determining \mathbf{a}_D , assumes the following maximum and minimum values, ρ_0, ρ_a in terms of the ellipse principal semi-axis lengths a, b .

$$\rho_0 = a^2/b, \quad \rho_a = b^2/a$$

An ellipse with $a/b = 2$, its pairs of circles of minimum and maximum curvature and the locus of curvature centres in all four quadrants are shown in Fig. 3. Typical normal and tangential directions are plotted for $x = a/2$. The normal segment is ρ_D long.

5.2 Computing the Curvature Centre Curve and Radius at Points on the Ellipse

The following short piece of code was used to compute the curvature centre curve for the 2 : 1 aspect ratio ellipse shown in Fig. 3. $D(XD, YD)$, $\rho_D = RH$, $|\mathbf{n}_D| = DN$, (XND, YND) are coordinates of points on the curvature centre curve.

```
100 A=2:B=1
110 FOR XD=0 TO SQR(A) STEP 0.1
```

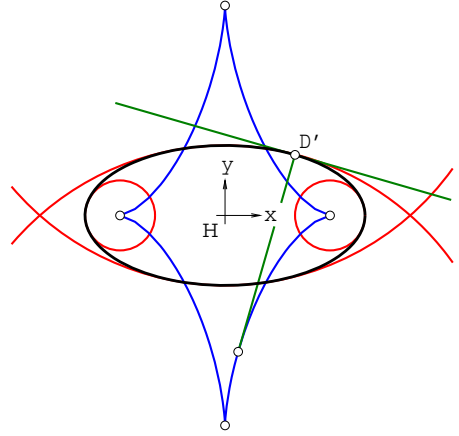


Figure 3: Curvature of Ellipse

```
120 YD=SQR ( ( A*B-B*XD^2 ) / A ) :
    RH=SQR ( ( A^2-A*XD^2+B*XD^2 ) ^ 3
        / ( B*A^4 ) )
130 DN=SQR ( ( B*XD ) ^ 2
    + ( A*YD ) ^ 2 ) :
    XND=XD-RH*B*XD/DN :
    YND=YD-RH*A*YD/DN
140 PRINT USING"#####.####";
    XND, YND, RH, YD
150 NEXT XD
160 STOP:END
```

6. CONCLUSION

The essence of this mechanism is a planar “slider-crank” wherein the “wrist-pin” moves on an ellipse rather than a straight line. Kinematic analysis is greatly simplified if a frame, that places this ellipse in standard form on a principal Cartesian

plane, is used. An explicit parametrization of the ellipse in terms of distance measured along its major axis, the x -axis in this case, provides for convenient computation of required variables, especially curvature radii, which facilitates the acceleration analysis. Last but not least we have here yet another example of an ellipse generating mechanism which may not have been previously reported as such.

ACKNOWLEDGMENTS

Partial support of this research by a Natural Sciences and Engineering Research Canada (NSERC) "Discovery" grant is gratefully acknowledged.

REFERENCES

- [1] J.E. Baker. A comparative survey of the Bennett-based, 6-revolute kinematic loops. *Mech. & Mach. Theory*, 28(1): 83–96, 1993.
- [2] G.T. Bennett. A new mechanism. *Engineering* 76: 777–778, 1903.
- [3] P. Dietmaier. *Einfach übergeschlossene Mechanismen mit Drehgelenken*. Habilitationsschrift, TU-Graz, 53–110, 1995.
- [4] M. Goldberg. New five-bar and six-bar linkages in three dimensions. *ASME Trans.* 65, 649–656, 1943.
- [5] C. Mavroidis and B. Roth. Analysis and synthesis of overconstrained mechanisms. *Mechanism Synthesis and Analysis*, ed. G.R. Pennock, ASME, New York, DE-70: 115–133, 1994.
- [6] C. Mavroidis and B. Roth. New and revised overconstrained mechanisms. *ASME J. Mech. Des.*, 117(1):75–82, 1995.
- [7] P.T. Sarrus. Note sur la transformation des mouvements rectilignes alternatifs, en mouvements circulaires et reciproquement. *Comtes. Rendus, Acad. Sci.*, Paris, 36: 1036–1038, 1853.
- [8] K.J. Waldron. A study of overconstrained linkages geometry by solution of closure equations, Part II, Four-bar linkages with lower pair joints other than screw joints. *Mech. & Mach. Theory* 8: 233–247, 1973.
- [9] K. Wohlhart. Merging two general Goldberg 5R linkages to obtain a new 6R space mechanism. *Mech. & Mach. Theory* 26(7): 659–668, 1991.
- [10] K. Wohlhart. On isomeric overconstrained space mechanisms. In *Proc. 8th IFToMM Wld. Congr.*, 1: 153–158, 1991.
- [11] K. Wohlhart. The two types of the orthogonal Bricard linkage. *Mech. & Mach. Theory*, 28(6): 809–817, 1993.
- [12] <<http://www.cim.mcgill.ca/~paul/LG.pdf>>.

ABOUT THE AUTHORS

1. Jesús Cervantes-Sánchez, in addition to kinematics, does research on computer vision and genetic algorithms. He is professor of mechanical engineering at Universidad de Guanajuato, México
2. Paul J. Zsombor-Murray's main interest is applying methods from geometry to problems in mechanics. He is associate professor of mechanical engineering (post-retirement) at McGill University in Montreal, Canada.

LEARN BY EATING – AN EASY WAY TO APPROACH GEOMETRICAL AND MATHEMATICAL ASPECTS

Elena MARCHETTI and Luisa ROSSI-COSTA

Politecnico di Milano, Italy

ABSTRACT: Geometry is one of the oldest branches of mathematics and accompanies people throughout her own experience. The knowledge of this subject, increasingly investigated since the time of Euclid, leads to discovering many crucial aspects, related not only to the development of mathematics but also to other scientific fields. It is very important motivate students, at any levels of learning, to improve the familiarity with this subject. Therefore, it is necessary to train pupils since primary school to recognize aspects of geometric elements even of everyday life.

The paper proposes to assume the food as vehicle of connection with geometry, taking inspiration from all that is presented on the table. That is not only what is concerned with ingredients of a recipe, but also what is necessary to prepare an elegant table.

Keywords: Applied Geometry in Architecture and Arts, Curves, Surfaces, Education.

1. INTRODUCTION

The forms of nature or manufactured by human creativity are important examples of geometric structures. If the mathematical concepts are related to elements of the nature (such as flowers and leaves, butterflies and different insects, shells, crystals,...) they become easier to be transferred to the students, but also to ordinary people. At the same time, the knowledge of geometric and analytical properties gives more opportunity to discover and identify mathematical elements around us as in Figure 1.



Figure 1: Tea pot (ats Design) ¹

¹ A. and T. Scarpa Design; by courtesy of San Lorenzo@ Silversmith, Milano. <http://www.sanlorenzossilver.it>

Therefore, thinking to involve more people on these topics and to give them the message that *Mathematics is everywhere*, we propose to connect food, that normally “brings well-being”, to geometry or to mathematical properties, frequently “related to anxiety”.

Since the founding of the Didactic Laboratory FDS - Department of Mathematics²- Politecnico di Milano, we have promoted mathematical projects related to nature, art, architecture, design, engineering and other fields. In particular we have developed the project *Imparare Mangiando*, that is *Learn by Eating*.

This approach has been tested out by us in different contexts: in occasion of commercial exhibitions on *Packaging for foods*³, in *Learning Weeks* for high-school students, and recently in EXPO DAYS⁴ promoted in Milan

² <http://fds.mate.polimi.it/>

³ http://fds.mate.polimi.it/index.php?arg=sperimentazione&id_pagina=175

⁴ http://fds.mate.polimi.it/index.php?arg=formazione&id_pagina=180

by EXPO 2015, whose theme is *Feeding the Planet, Energy for Life*.

In the traditional Italian food we can observe many forms: just think of the different kinds of pasta, suitable to be described through mathematical tools and reproduced by computer.

It is a good opportunity to study, on a laid table, elements involving mathematical and geometrical aspects.

We have some aims to be pursued:

1. the study of curves and surfaces, underlining their geometrical and differential peculiarities,
2. the study of symmetries in friezes and rosettes,
3. the introduction of algorithms, proportions and self-similarity.

Geometry and Mathematical Analysis help us in composing a gorgeous and rich menu: different starters prepared with geometric fantasy, pasta in all shapes and sizes, seafood with their spirals served with the roman cauliflower, a vegetable fractal. *Dulcis in fundo*, a collection of cakes and pastries decorated by friezes and rosettes. To conclude a strong Italian coffee, served with sugar lumps, not only of cubic form but also of other Platonic solids realized by a new 3D printer in Figure 2.



Figure 2: Platonic solids⁵

2. MATHEMATICAL TOOLS

Curves and surfaces are the boundary elements of all objects: it is very important to give the

students a mathematical way to describe them.

Among different possibilities we choose the parametric form, one of the most adaptable and efficient way to represent a geometric *locus*, both graphically and analytically [1]. More in details, in the 3D Cartesian orthogonal space $Oxyz$, a *regular curve* is represented by the column vector

$$\mathbf{v}(t) = [x(t), y(t), z(t)]^T$$

depending on the *real parameter*

$$t \in I = [a, b] \subset \mathbb{R}, \quad \mathbf{v}(t) \in C^1(I) \quad \text{and} \quad |\mathbf{v}'(t)| \neq 0,$$

that is the usual *conditions of regularity* are satisfied.

Equally a *regular surface* is represented by the column vector

$$\mathbf{w}(t, u) = [x(t, u), y(t, u), z(t, u)]^T$$

depending on *two real parameters*

$$(t, u) \in A = ([a, b] \times [c, d]) \subset \mathbb{R}^2$$

$$\mathbf{w}(t, u) \in C^1(A) \quad \text{and} \quad |\mathbf{w}_t \times \mathbf{w}_u| \neq 0.$$

It is also very common to observe the different aspects of symmetry, focusing on 2D and 3D geometrical transformations. Therefore we need to introduce their mathematical representation through vectors and matrices. By the notation $\mathbf{w} = \mathbf{A}\mathbf{v}$, we indicate that the vector \mathbf{v} is transformed in \mathbf{w} by the matrix \mathbf{A} , where the column vectors \mathbf{v} and \mathbf{w} have two or three components and consequently \mathbf{A} is respectively a (2, 2) or (3, 3) square matrix [2].

In the following we propose an elementary example: a circle of radius $r=1$ belonging to the Oxy plane and centered in O can be represented by the two components column vector $\mathbf{v}(t) = [\cos t, \sin t]^T, t \in [0, 2\pi]$. The

$$\text{product} \quad \mathbf{w}(t) = \mathbf{A}\mathbf{v}(t) \quad \text{with} \quad \mathbf{A} = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$$

(a, b positive real numbers) leads to the vector

$$\mathbf{w}(t) = [a \cos t, b \sin t]^T, t \in [0, 2\pi]$$

representing the ellipse with semi-axes a, b and centered in the origin O .

It is also very important introduce young students to algorithms. An algorithm is expressed by a finite list of well-defined

⁵ By courtesy of +LAB, Director Marinella Levi [5], Dipartimento di Chimica, Materiali e Ingegneria Chimica "Giulio Natta", Politecnico di Milano <http://www.piulab.it>

instructions. It is used to reach a numerical result, or to develop an automated reasoning. Frequently a numerical result could be obtained only as a limit of an infinite number of recursive steps but the algorithm can provide (for any fixed number n of steps) the n -th approximation of the result.

The idea of the finite number of steps belonging to an infinite sequence can describe a fractal, an auto-similar geometric structure.

2.1 Ruled surfaces

Planes, cylinders and cones are ruled surfaces, that is for each their point passes at least one straight line.

The classical *rigatone* (or *maccherone*) is a portion of a circular cylinder between two planes orthogonal to its axis (Figure 3)⁶.

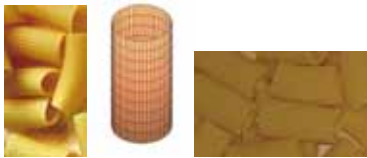


Figure 3: *Maccheroni* and circular cylinder

We generate the so called *penna* by cutting the circular cylinder with two parallel planes inclined with respect to its axis. In this case the boundary section is an ellipse (Figure 4).



Figure 4: *Penne* and elliptical sections

Usually we call cylinder any surface generated by moving a straight line, in a

parallel way to a fixed one (called *generatrix*), along a curve called *directrix*.

Other cylindrical *pastas* can be generated changing the *directrix*: e.g. *astroid*, *lemniscate*, *epicycloid* (Figure 5).

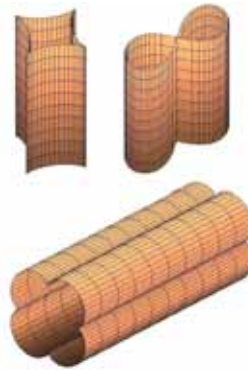


Figure 5: New cylindrical *pasta*

2.2 Helices

The pasta shape, called *fusillo*, is geometrically related to the cylindrical helix, a regular curve belonging to a cylindrical surface. This helix can be mathematically described by the following vector

$$\mathbf{v}(t) = [r \cos t, \quad r \sin t, \quad ht]^T$$

where t is a real parameter, r the radius of the cylinder and $h/2\pi$ indicates the elevation of a point of the curve after a complete turn, that is after the rotation of the angle of 2π .

It can be noted that

$$\mathbf{v}(t) = [t \cos t, \quad t \sin t, \quad ht]^T \quad (1)$$

describes an helix belonging to a circular cone.

From the cylindrical helix we can obtain a ruled surface, called *helicoid*, joining with a straight segment, the points of the vertical axis of the cylinder with the points of the helix.

Some kinds of pasta can be considered as portions of helicoids. Frequently the *fusillo* is also composed by two helicoids, one rotated with respect the other of half a turn around its axis (Figure 6).

⁶ All the graphic illustrations are generated by MATLAB®, starting from parametric equations.



Figure 6: Helix, helicoid and *fusilli*

The *torus* is the surface generated by turning a circle of radius r' around a circle of radius r ($r' < r$). The two circles belong to orthogonal planes. A portion of a *torus* is the geometrical basic form of *pipa* one of the most popular kind of *pasta corta* (short pasta) (Figure 7).

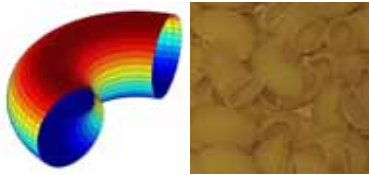


Figure 7: Torus and *pipe*

The translation of a circle of radius r' around a cylindrical helix ($r' < r/2$) characterizes another shape of pasta, called *cellentano*. Moreover sliding a circle along a sinusoidal curve we generate a new form: we name it *caterpillar* (Figure 8).

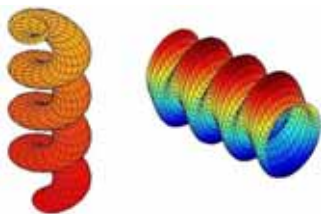


Figure 8: *Cellentano* and *caterpillar*

2.3 The *Pandoro*

The *Pandoro* is a traditional Italian cake, usually produced at Christmas time (Figure 9).



Figure 9: *Pandoro*

It has an interesting geometrical form: you can identify segments of straight line and curved lines (e.g. arcs of circle and of ellipse), circular plane portions, portions of cone or of a torus with elliptical section. We used all of these elements in the virtual reconstruction (Figure 10).

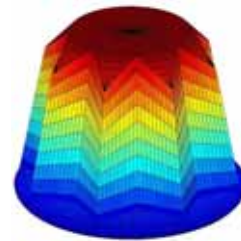


Figure 10: *Pandoro* virtual form

We can mention the parametric equations of some elements used in the reconstruction of this tasty *Christmas cake*, initially produced in Verona:

- the straight edges, as straight line segment between two points $A(x_0, y_0, z_0)$ and $B(x_1, y_1, z_1)$, can be represented by the vector

$$\mathbf{v}(t) = [(1-t)x_0 + tx_1, (1-t)y_0 + ty_1, (1-t)z_0 + tz_1]^T$$

where $0 \leq t \leq 1$;

- an ellipse on the vertical plane Oxz , with semi-axes of lengths a and b and center in $O(0,0,0)$, is described by the vector

$$\mathbf{u}(t) = [a \cos t, 0, b \sin t]^T \text{ where } 0 \leq t \leq 2\pi$$

(exchanging the second and the third

component of the vector the curve belongs to the Oxy plane).

The boundary surface of the *Pandoro* is composed by plane portions and curved surfaces. We mention only the equation of the torus surface with elliptical section, used in the reconstruction of the cake's base $\mathbf{w}(u,v)=[(r+a\cos u)\cos v, (r+a\cos u)\sin v, b\sin u]^T$ where a and b are the semi-axes of the ellipse, r ($r > a > b$) is the radius of the rotation circle at the base of the *Pandoro*, and $0 \leq v \leq 2\pi$, $-\pi/2 \leq u \leq \pi/2$.

The most of the horizontal sections of the *Pandoro* are octagonal stars. In this way we can outline symmetry axes in the sections and planes of symmetry in the whole cake (Figure 11).

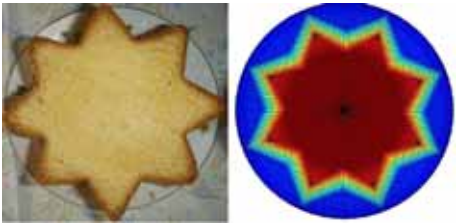


Figure 11: Star sections

3. A MATHEMATICAL MENU

In the previous section we started describing examples of food as mathematical objects. In this context we present a rich Italian menu composed of tasty foods connected to mathematical elements.

For starters we can choose a wide selection of canapés with symmetrical decorations, *tarallucci* of toroidal form, salami's elliptical slices and chips, shaped like hyperbolic paraboloid (in Figure 12 also its level curves).

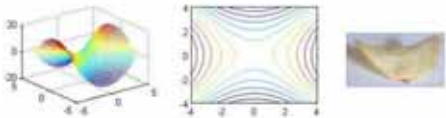


Figure 12: Hyperbolic paraboloid and chips

As a first course we propose *fusilli* with *broccoli*. This kind of vegetable belongs to the family of cruciferous; their inflorescences are composed of parts similar to the whole, that is they are the same at different scales as the fractals. In nature, one of the best examples of fractal is the Roman cauliflower (*brassica oleracea botrytis*) belonging to the same family of *broccoli*. In this vegetable we can distinguish not only self-similar parts but also conical structures where conical spirals are evident (Figure 13). It looks like an amazing sculpture!



Figure 13: Roman cauliflower

The boiled Roman cauliflower will be served with the main course, consisting of slices of meat, rolled with ham and spinach. Often this type of slices lead us to think of the Archimedes spiral (Figure 14).

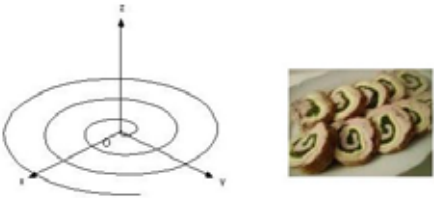


Figure 14: Archimedes spiral

This spiral can be represented by the vector $\mathbf{u}(t)=[t\cos t, t\sin t, 0]^T$ corresponding to vector (1) with the third component equal to zero. The Archimedes spiral is the projection of the conical spiral on the horizontal plane.

As dessert it is possible to prepare some cakes and tarts decorated with creativity and

fantasy, following the same symmetry of the cyclic or dihedral rosettes (Figure 15).

Why don't draw inspiration from the rose windows of the Milan Cathedral? [4]



Figure 15: Cyclic and Dihedral rosettes in cakes and in the Milan Duomo

In a cyclic rosette a non-symmetric basic motif is repeated n -times turning around a center with angle of rotation equal to $2\pi/n$.

In a dihedral rosette the basic motif has an axial symmetry, consequently the dihedral rosette presents n symmetry axes, passing through the center of rotation.

And now we propose a toast with a sparkling white wine, like *Prosecco*, served in elegant Murano glasses⁷ realized with mathematical forms and decorative curves. (Figure 16).



Figure 16: Glasses of Moretti collection

Finally a tea or a strong Italian coffee can be served in artistic conical cups, designed and decorated by the famous Italian architect Giò Ponti (Figure 17).



Figure 17: Coffee and tea service⁸

Sugar lumps with the shape of platonic solids, prepared by a 3D printer, will sweeten the drinks.

Mathematical aspects can come out also by a laid table: a beautiful embroidered tablecloth, some decorated dishes, cristal glasses... can make evident patterns and forms in 2D and 3D. The round tablecloth and the dish⁹ in Figure 18 have the same kind of decoration, from a mathematical point of view. Both are dihedral rosettes with the same number of symmetry axes.



Figure 18: Dihedral rosettes

Fruits and flowers can be placed in artistic silver bowls of particular geometrical forms (Figure 19). The bottom of the bowls is a polygonal silver slab and the upper edge is again a regular polygon with twice the sides in relation to the base.

The connection of the sides of the two polygons is realized in an elegant and smooth way. The lateral surfaces of the bowls remind us of the minimal surfaces [3].

The mathematical investigation in this subject is not so easy, but an amazing way to familiarize the students on minimal surfaces is

⁷ Carlo and Giovanni Moretti - Murano

⁸ Giò Ponti Collection by Richard-Ginori around 1930

⁹ Giò Ponti Collection by Richard-Ginori-Gucci, 2014

to play with soap bubbles.



Figure 19: ats Design¹⁰

Finally we observe that, preparing a laid table and following a recipe, we do a sequence of steps like in an algorithm. In particular, when cooking a meal, all the ingredients used must be proportional to the number of guests. A logical-mathematical reasoning must always be developed.

4. CONCLUSIONS

One of the main goal of the FDS Lab is to connect mathematics with other cultural fields, not only scientific. We believe that all the students, especially for their university studies, have to be prepared in discovering the different aspects of culture.

We have presented this work so that students become more curious and confident towards mathematics. They have many occasions to read *mathematics everywhere*, even during a relaxing meal.

Food, a primary need, is the crucial topic of EXPO 2015 in Milan, and this will be a good occasion to discover more about it.

REFERENCES

- [1] R. Adams and C. Essex. *Calculus: A Complete Course*, 8th ed. Pearson Education Canada, Toronto, 2014.
- [2] R. Betti, E. Marchetti and L. Rossi Costa (Editors). *Simmetria: una scoperta matema-*

tica, 2nd ed. Polipress, Politecnico di Milano, 2012.

- [3] M. Emmer. *Minimal Surfaces and Architecture: New Forms*. Nexus Netw J, 15 (2): 227-239, 2013.
- [4] E. Marchetti and L. Rossi Costa. What geometry in Milan Cathedral?. *Nexus VI – Architecture and Mathematics*, Kim Williams Books, 63-76, 2006.
- [5] F. Ostuzzi, V. Rognoli and M. Levi. +TUO: 3D Printing as tool for co-design with and for users with Rheumatic Diseases. *A Matter of Design: Making Society through Science and Technology*. 5th STS Italia Conference – Milan, 12-14 June 2014 (in press)

ABOUT THE AUTHORS

1. Elena Marchetti, Associate Professor, since 1988 has taught Mathematics in courses for architecture students at the School of Architecture of the Politecnico di Milano. She has produced numerous publications in Italian and international scientific journals in Numerical Analysis. She has published many papers about the applications of Mathematics in Architecture, Design and Arts. She collaborated to books dedicated to this topic, with multimedia support packages.

2. Luisa Rossi Costa, Associate Professor of Mathematical Analysis in the Engineering Schools of the Politecnico di Milano, developed researches in Numerical and Functional Analysis. She contributed to create e-learning platforms and since 1998, she works in formation projects for Students and Teachers of High Schools. She wrote numerous papers on teaching methods and on connections between Art, Design, Architecture and Mathematics. She has been Editor with colleagues of a book on symmetry, with a DVD support.

¹⁰ A. and T. Scarpa Design; by courtesy of San Lorenzo@ Silversmith, Milano.
<http://www.sanlorenzsilver.it>

LINE-INVERSION IN THE QUASI-HYPERBOLIC PLANE

Helena HALAS, Nikolina KOVAČEVIĆ and Ana SLIEPČEVIĆ

University of Zagreb, Croatia

ABSTRACT: In this paper an involutive quadratic transformation called the line-inversion is introduced in the quasi-hyperbolic plane. The proper images of the pencils of lines and the 2nd class curves obtained by the line-inversion using different types of the transformation will be analysed.

Keywords: quasi-hyperbolic plane, line-inversion, equiform curve

1. INTRODUCTION

The quasi-hyperbolic plane \mathbb{QH}_2 is one of nine Cayley-Klein projective metric planes. It is a real projective plane $\mathcal{P}_2 = \{\mathcal{P}, \mathcal{L}, \mathbf{I}\}$ where the metric is induced by a real degenerated conic consisting of a triple $\{F, f_1, f_2\}$, where f_1 and f_2 are two real lines intersecting at the point F , [4–7]. This triple determines the absolute figure $\mathcal{F}_{\mathbb{QH}}$ of the quasi-hyperbolic plane. The point F is called the *absolute point* and lines f_1, f_2 are called the *absolute lines*.

1.1 Analytical background

We consider the plane \mathbb{QH}_2 embedded in its projective extension, the real complexified projective plane $\mathcal{P}_2 = PG(2, \mathbb{R}) \subset PG(2, \mathbb{C})$ where points and lines are the equivalence classes of the set $\mathbb{R}^3 \setminus \{(0, 0, 0)\}$ and will be denoted by the same letters as their homogeneous 3-vectors, i.e. the point $\mathbf{x} = (x_0 : x_1 : x_2)^1 \sim \lambda \mathbf{x}$ and the line $\bar{u} = [u_0 : u_1 : u_2] \sim \mu \bar{u}$, $\lambda, \mu \in \mathbb{R} \setminus \{0\}$.

By introducing a canonical scalar product the condition of incidence is given with

$$\mathbf{x} \cdot \bar{u} = x_0 u_0 + x_1 u_1 + x_2 u_2 = 0.$$

Furthermore, an operation of vector product on $\mathbb{R}^3 \times \mathbb{R}^3$ serves as a basis for introducing the \wedge -product:

$$\begin{aligned} (\alpha_0, \alpha_1, \alpha_2) \wedge (\beta_0, \beta_1, \beta_2) &:= \\ (\alpha_1 \beta_2 - \alpha_2 \beta_1, \alpha_2 \beta_0 - \alpha_0 \beta_2, \alpha_0 \beta_1 - \alpha_1 \beta_0). \end{aligned} \quad (1)$$

¹When using matrix notation we regard \mathbf{x} as a column vector, even if we write \mathbf{x} for simplicity.

It can be used in two variants, $\wedge : \mathcal{P}_2 \times \mathcal{P}_2 \mapsto \mathcal{P}_2^*$, or $\wedge : \mathcal{P}_2^* \times \mathcal{P}_2^* \mapsto \mathcal{P}_2$, where \mathcal{P}_2^* is the dual plane of \mathcal{P}_2 . Also, based on a simple calculations, the validity of the following equations can be shown:

Lemma 1. *Let a real projective plane $\mathcal{P}_2 = \{\mathcal{P}, \mathcal{L}, \mathbf{I}\}$ be given. For $\mathbf{x}, \mathbf{y} \in \mathcal{P}$ and $\bar{u}, \bar{v} \in \mathcal{L}$, we have²:*

$$1) (\mathbf{x} \wedge \mathbf{y}) \wedge \bar{u} = (\mathbf{x} \cdot \bar{u}) \mathbf{y} - (\mathbf{y} \cdot \bar{u}) \mathbf{x} := \mathbf{z},$$

$$2) (\bar{u} \wedge \bar{v}) \wedge \mathbf{x} = (\mathbf{x} \cdot \bar{u}) \bar{v} - (\mathbf{x} \cdot \bar{v}) \bar{u} := \bar{w}.$$

For the 2nd class curve ζ the quadratic form $\omega_C : \mathbb{R}^3 \rightarrow \mathbb{R}$ and the associated symmetric bilinear form $f_C : \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}$ are given with

$$\omega_C(\bar{u}) = \sum_{i,j=0}^2 c_{ij} u_i u_j = 0, \quad C = (c_{ij}), \quad c_{ij} \in \mathbb{R}, \quad (2)$$

$$f_C(\bar{u}, \bar{v}) = \frac{1}{2} (\omega_C(\bar{u} + \bar{v}) - \omega_C(\bar{u}) - \omega_C(\bar{v})) = 0.$$

Hence, it follows $\omega_C(\bar{u}) = f_C(\bar{u}, \bar{u})$ and f_C is called the polar form of ω_C with respect to ζ . Furthermore, two lines \bar{u}, \bar{v} are called *conjugate lines* with respect to ζ , iff $f_C(\bar{u}, \bar{v}) = 0$. Also, the 2nd class curve ζ induces a *polarity* given with $\pi_\zeta : \mathcal{L} \mapsto \mathcal{P}$ and $\pi_\zeta^{-1} : \mathcal{P} \mapsto \mathcal{L}$

$$\mathbf{u} := \pi_\zeta(\bar{u}) = C\bar{u}, \quad (3)$$

$$\bar{v} := \pi_\zeta^{-1}(\mathbf{v}) = \mathbf{v}C^{-1}. \quad (4)$$

²Although the vector product is an anticommutative operation, the use of the homogeneous coordinates justifies that the \wedge -product operation is commutative in \mathcal{P}_2 .

In order to parameterize the plane \mathbb{QH}_2 , we have placed the coordinate simplex of the embedded plane \mathcal{P}_2 in a such way that the absolute figure $\mathcal{F}_{\mathbb{QH}}$ is determined by

$$\bar{f}_1 \sim [0:1:1], \quad \bar{f}_2 \sim [0:-1:1], \quad \mathbf{f} \sim (1:0:0).$$

The lines through \mathbf{f} are called the *isotropic lines*, and the points incidental with the line \bar{f}_1 or \bar{f}_2 are called *isotropic points*.

In the sense of the Erlangen program, for the fundamental group of transformations in \mathbb{QH}_2 we use the *4-parameter general quasi-hyperbolic group of similarities* \mathfrak{G}_4 , [4]. It keeps the absolute figure $\mathcal{F}_{\mathbb{QH}}$ invariant, and the transformations are of the form $[u_0 : u_1 : u_2] \mapsto [\alpha_0 u_0 : \alpha_1 u_0 + \alpha_2 u_1 + \alpha_3 u_2 : \alpha_4 u_0 \pm \alpha_3 u_1 \pm \alpha_2 u_2]$, $\alpha_i \in \mathbb{R}$, $i = \{0 \dots 4\}$, $\pm \alpha_2^2 \pm \alpha_3^2 \neq 0$.

2. THE 2ND CLASS CURVES IN \mathbb{QH}_2

Further in text we will refer to the 2nd class curves as line-conics to point out that we are observing them as line envelopes, although it is often more convenient to draw the figures as point objects as we are used to.

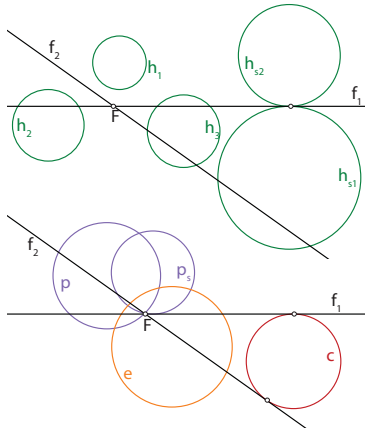


Figure 1: Types of line-conics in the projective model of \mathbb{QH}_2

In accordance to the group \mathfrak{G}_4 , non-degenerate line-conics in \mathbb{QH}_2 can be classified

into nine types, [5]. A line-conic that has no, one (double) or two real isotropic lines is called an *ellipse* (e), *parabola* (p) or *hyperbola* (h_1, h_2, h_3), accordingly. If one of the isotropic line of the line-conic coincides with an absolute line it is called *special parabola* (p_s) or a *special hyperbola* (h_{s1}, h_{s2}), and if both isotropic lines coincide with the absolute lines it is called a *circle* (c) (see Fig. 1).

In the chosen projective frame every line-conic ζ can be represented by (2) and the coordinates of its isotropic lines $\bar{u} \sim [0 : u_1 : u_2]$ satisfy the equality

$$c_{11}u_1^2 + c_{22}u_2^2 + 2c_{12}u_1u_2 = 0. \quad (5)$$

Some short calculation lead us to the conclusions that ζ is:

$$\text{an ellipse iff } c_{12}^2 - c_{11}c_{22} < 0, \quad (6)$$

$$\text{a parabola iff } c_{12}^2 - c_{11}c_{22} = 0, \quad (7)$$

$$\text{a hyperbola iff } c_{12}^2 - c_{11}c_{22} > 0, \quad (8)$$

$$\text{a special parabola iff } c_{11} = c_{22} = \pm c_{12}, \quad (9)$$

$$\text{a special hyperbola iff } c_{11} + c_{22} \pm 2c_{12} = 0, \quad (10)$$

$$\text{a circle iff } c_{12} = 0, c_{11} = -c_{22}. \quad (11)$$

A *central line* \bar{c} of a line-conic ζ is the polar line of the absolute point i.e. $\bar{c} = \pi_{\zeta}^{-1}(\mathbf{f})$. The line-conics with the same central line are called *concentral* line-conics.

Some further results on basic notions and line-conics in \mathbb{QH}_2 may be found in [5].

2.1 Equiform line-conics

The orthogonality in \mathbb{QH}_2 is ruled by $\mathcal{F}_{\mathbb{QH}}$, i.e. the metric is induced by the hyperbolic involution on the pencil (\mathbf{f}) whose double lines are \bar{f}_1, \bar{f}_2 , [4]. It's called the *absolute involution* of \mathbb{QH}_2 , denoted by $\mathcal{I}_{\mathbb{QH}}$, and its explicit form is given with

$$\mathcal{I}_{\mathbb{QH}} : [0 : u_1 : u_2] \mapsto [0 : u_2 : u_1].$$

Each line-conic ζ given by (2), except parabolas, induces the involution ϕ_{ζ} on the pencil (\mathbf{f})

where the double lines \bar{q}_1, \bar{q}_2 are the isotropic lines of ζ . The notion of equiformally specified line-conics is introduced in \mathbb{QH}_2 based on the mutual position of lines $\bar{q}_1, \bar{q}_2, \bar{f}_1$ and \bar{f}_2 . Thus, we distinguish two classes of such line-conics.

The *first class* is formed by line-conics whose isotropic lines \bar{q}_1, \bar{q}_2 coincide with the absolute lines \bar{f}_1, \bar{f}_2 , i.e. $\phi_C = \mathcal{I}_{\mathbb{QH}}$. Those line-conics are circles whose coefficients satisfy the condition (11) and their central lines are $\bar{c} \sim [c_{11} : -c_{01} : c_{02}]$. By using the translation from \mathfrak{G}_4 given with $[u_0 : u_1 : u_2] \mapsto [u_0 : -\frac{c_{01}}{c_{11}}u_0 + u_1 : \frac{c_{02}}{c_{11}}u_0 + u_2]$ when $c_{11} \neq 0$, the circle can be transformed into the form

$$au_0^2 + u_1^2 - u_2^2 = 0, \quad (12)$$

$$\bar{c} \sim [1 : 0 : 0], \quad a \in \mathbb{R} \setminus \{0\}.$$

Also by choosing the suitable translation and coefficients these circles can be given with the equation

$$u_0^2 + \frac{a^2}{2}u_1^2 - \frac{a^2}{2}u_2^2 + 2au_0u_1 = 0, \quad (13)$$

$$\bar{c} \sim [-a/2 : 1 : 0], \quad a \in \mathbb{R} \setminus \{0\}.$$

In the *second class* are the line-conics for which the isotropic lines are in harmonic relation with the absolute lines i.e. $H(\bar{f}_1\bar{f}_2, \bar{q}_1\bar{q}_2) = -1$. If the double lines of the involution ϕ_C of line-conic ζ are given with $\bar{q}_1 \sim [0 : q_{11} : q_{12}]$ and $\bar{q}_2 \sim [0 : q_{21} : q_{22}]$, then from the harmonic relation follows

$$q_{12}q_{22} - q_{11}q_{21} = 0. \quad (14)$$

On the other hand, the lines \bar{q}_1 and \bar{q}_2 are the associated lines of ϕ_C , therefore from (5) and (14) for the coefficients of ζ in (2) it follows

$$c_{11} = c_{22}.$$

By using the translation from \mathfrak{G}_4 $[u_0 : u_1 : u_2] \mapsto [u_0 : \frac{c_{02}c_{12} - c_{01}c_{11}}{c_{11}^2 - c_{12}^2}u_0 + u_1 : \frac{c_{01}c_{12} - c_{02}c_{11}}{c_{11}^2 - c_{12}^2}u_0 + u_2]$ since $(c_{11}^2 - c_{12}^2) \neq 0$, those line-conics can be transformed into the form

$$\hat{c}_{00}u_0^2 + \hat{c}_{11}u_1^2 + \hat{c}_{11}u_2^2 + 2\hat{c}_{12}u_1u_2 = 0, \quad (15)$$

$$\hat{c}_{00}, \hat{c}_{11}, \hat{c}_{12} \in \mathbb{R} \setminus \{0\}.$$

Depending whether the induced involution ϕ_C is hyperbolic or elliptic i.e. depending on the reality of the discriminant $\Delta = c_{12}^2 - c_{11}c_{22}$ in (5), the line-conic ζ in the second class of equiformally specified line-conics is a hyperbola or an ellipse, [2], [5]. Hence, for $\hat{c}_{11} = 0$ in (15) it is the hyperbola given with the equation

$$au_0^2 + 2u_1u_2 = 0 \quad (16)$$

$$\bar{c} \sim [1 : 0 : 0], \quad a \in \mathbb{R} \setminus \{0\}.$$

and for $\hat{c}_{12} = 0$ the ellipse given with the equation

$$au_0^2 - u_1^2 - u_2^2 = 0, \quad (17)$$

$$\bar{c} \sim [1 : 0 : 0], \quad a \in \mathbb{R} \setminus \{0\}.$$

Also by choosing the suitable translation and coefficients the aforementioned hyperbolas can be given with the equation

$$u_0^2 + 2au_0u_1 + 2au_0u_2 + 2a^2u_1u_2 = 0, \quad (18)$$

$$\bar{c} \sim [-a : 1 : 1], \quad a \in \mathbb{R} \setminus \{0\}.$$

To summarize these results we introduce the following definition of the special class of the line-conics in \mathbb{QH}_2 :

Definition 1. A line-conic ζ for which the induced involution ϕ_C commutes with the absolute involution $\mathcal{I}_{\mathbb{QH}}$ is called an *equiform line-conic*.

According to some further results on basic notions in \mathbb{QH}_2 (see [5]), it follows from (4)

Proposition 1. The poles of the absolute lines with respect to the equiform line-conic are isotropic points.

3. LINE-INVERSION

Let a line \bar{p} and a non-degenerate line-conic κ with its matrix equations be given

$$\kappa : \bar{u}^T \mathbf{K} \bar{u} = 0, \quad K = (k_{ij}),$$

$$\bar{p} : \bar{p} \cdot \mathbf{x} = 0, \quad \bar{p} \sim [p_0 : p_1 : p_2],$$

in the projective plane \mathcal{P}_2 .

Definition 2. A line-inversion $\sigma_{(\bar{p}, \kappa)}$ with respect to the line \bar{p} and the line-conic κ is a transformation where the corresponding lines are conjugate with respect to κ and are concurrent at a point on \bar{p} .

The line \bar{p} and the line-conic κ are called the *polar line* and *fundamental line-conic* of the line-inversion $\sigma_{(\bar{p}, \kappa)}$. The polar line \bar{p} and the double lines \bar{p}_1, \bar{p}_2 of the involution determined in the pencil of lines $(\pi_\kappa(\bar{p}))$ are called *fundamental lines* and form a *fundamental trilateral* of the given birational involutive mapping $\sigma_{(\bar{p}, \kappa)}$, [1–3]. The vertices of the fundamental trilateral are the *fundamental points* $\pi_\kappa(\bar{p})$, $\pi_\kappa(\bar{p}_1)$, $\pi_\kappa(\bar{p}_2)$, denoted by \mathbf{p} , \mathbf{p}_1 and \mathbf{p}_2 respectively. Besides the lines of the pencils (\mathbf{p}) , (\mathbf{p}_1) , (\mathbf{p}_2) the transformation $\sigma_{(\bar{p}, \kappa)}$ is an one-to-one mapping.

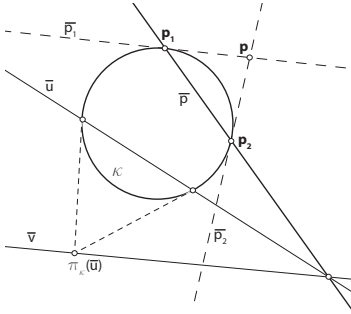


Figure 2: Image \bar{v} of the line \bar{u} with respect to $\sigma_{(\bar{p}, \kappa)}$

In general, to each line \bar{u} we associate its image line \bar{v}

$$\bar{v} \sim \sigma_{(\bar{p}, \kappa)}(\bar{u}) = (\bar{u} \wedge \bar{p}) \wedge \pi_\kappa(\bar{u}) \quad (19)$$

where by Lemma 1 it follows

$$\sigma_{(\bar{p}, \kappa)}(\bar{u}) = \omega_\kappa(\bar{u}) \cdot \bar{p} - f_\kappa(\bar{u}, \bar{p}) \cdot \bar{u}. \quad (20)$$

Therefore if

$$\bar{u} \in (\mathbf{p}) \rightarrow f_\kappa(\bar{u}, \bar{p}) = 0, \text{ hence } \bar{v} \sim \bar{p}, \quad (21)$$

$$\begin{aligned} \bar{u} \in (\mathbf{p}_i) &\rightarrow \bar{v} \sim (\bar{u} \wedge (\mathbf{p}_1 \wedge \mathbf{p}_2) \wedge \mathbf{u}) \\ &\sim (\mathbf{p}_i \wedge \pi_\kappa(\bar{u})) \sim \bar{p}_i, \quad i \in \{1, 2\}, \end{aligned} \quad (22)$$

the lines of the pencils (\mathbf{p}) , (\mathbf{p}_1) , (\mathbf{p}_2) are mapped to the associated fundamental lines and the lines of κ are invariant under $\sigma_{(\bar{p}, \kappa)}$ because

$$\bar{u} \in \kappa \rightarrow \omega_\kappa(\bar{u}) = 0, \text{ hence } \bar{v} \sim \bar{u}. \quad (23)$$

Furthermore, since the fundamental lines are the singular lines of $\sigma_{(\bar{p}, \kappa)}$ their images are defined to be the corresponding pencils of fundamental points.

Generally, $\sigma_{(\bar{p}, \kappa)}$ maps a curve k^n of class n into the curve k_κ^{2n} of class $2n$. The generating curve k^n has n lines passing through the fundamental points, thus by (21) and (22) the fundamental lines are n -fold lines of the curve k_κ^{2n} . If the generating curve k^n contains one fundamental line as r -fold line, then k_κ^{2n} splits into the corresponding pencil of lines with multiplicity r and the curve of class $2n - r$, [1, 2].

The image of an arbitrary pencil $\mathbf{q}\bar{u} = 0$ with respect to $\sigma_{(\bar{p}, \kappa)}$ is the line-conic k_κ^2

$$k_\kappa^2 : \omega_\kappa(\bar{u}) \cdot \mathbf{q}\bar{p} - f_\kappa(\bar{u}, \bar{p}) \cdot \mathbf{q}\bar{u} = 0,$$

and if

- $\mathbf{q} \in \bar{p} \rightarrow (\mathbf{q}) \cap \kappa = \{\bar{t}_1, \bar{t}_2\}$, by (23) k_κ^2 degenerates into (\mathbf{p}) and (\mathbf{q}) ,
- $\mathbf{q} \in \bar{p}_i, i \in \{1, 2\} \rightarrow (\mathbf{q}) \cap \kappa = \{\bar{p}_i, \bar{t}\}$, by (23) k_κ^2 degenerates into (\mathbf{p}_i) and (\mathbf{q}') where $\mathbf{q}' = \bar{t} \cap \bar{p}_j, j \in \{1, 2\}, j \neq i$.

Theorem 1. A line-inversion $\sigma_{(\bar{p}, \kappa)}$ in \mathcal{P}_2 maps the pencil (\mathbf{q}) , not containing the fundamental lines, into the line-conic k_κ^2 containing all fundamental lines. It also contains lines $\bar{q}, \bar{t}_1, \bar{t}_2$, where \bar{q} is the polar line of the point \mathbf{q} with respect to κ , and \bar{t}_1, \bar{t}_2 are common lines of κ and (\mathbf{q}) .

Proof: Let be $\bar{a}, \bar{b}, \bar{c} \in (\mathbf{q})$ such that $\mathbf{p} \in \bar{a}$, $\mathbf{p}_1 \in \bar{b}$, $\mathbf{p}_2 \in \bar{c}$. Therefore by (21) and (22) the fundamental lines are the lines of the image of the pencil (\mathbf{q}) . Furthermore, the lines \bar{t}_1 and \bar{t}_2 common to (\mathbf{q}) and κ are by (23) invariant to $\sigma_{(\bar{p}, \kappa)}$, hence they are also the lines of the image (see fig. 3). For the lines $\bar{u}_\lambda \in (\mathbf{q})$, $\lambda \in \mathbb{R}$, it follows $\pi_\kappa(\bar{u}_\lambda) \in \bar{q}$ where $\bar{q} = \pi_\kappa^{-1}(\mathbf{q})$. The image

of the pencil (\mathbf{q}) is the result of two projective point ranges on the lines \bar{p} and \bar{q} , hence it is the line-conic containing the line \bar{q} as well. \square

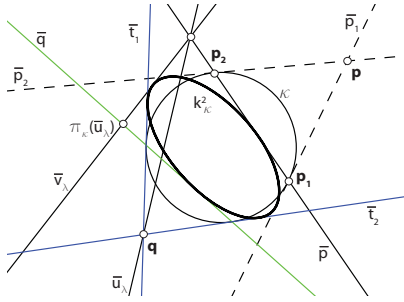


Figure 3: Image k_K^2 of the pencil (\mathbf{q}) with respect to $\sigma_{(\bar{p}, \kappa)}$

The image of an arbitrary line-conic q^2 is a curve k_K^4 of the 4th class and analogous to the previous observations it will degenerate in the following cases

- i) $\bar{p} \in q^2$ or $\bar{p}_i \in q^2, i \in \{1, 2\}$
 $\sigma_{(\bar{p}, \kappa)} : q^2 \rightarrow (\mathbf{p}) \cup k_K^3$ or
 $\sigma_{(\bar{p}, \kappa)} : q^2 \rightarrow (\mathbf{p}_1) \cup k_K^3,$
- ii) $\bar{p}, \bar{p}_i \in q^2, i \in \{1, 2\}$ or $\bar{p}_1, \bar{p}_2 \in q^2$
 $\sigma_{(\bar{p}, \kappa)} : q^2 \rightarrow (\mathbf{p}) \cup (\mathbf{p}_1) \cup k_K^2$ or
 $\sigma_{(\bar{p}, \kappa)} : q^2 \rightarrow (\mathbf{p}_1) \cup (\mathbf{p}_2) \cup k_K^2,$
- iii) $\bar{p}, \bar{p}_1, \bar{p}_2 \in q^2$
 $\sigma_{(\bar{p}, \kappa)} : q^2 \rightarrow (\mathbf{p}) \cup (\mathbf{p}_1) \cup (\mathbf{p}_2) \cup k_K^1.$

The curves $k_K^\alpha, \alpha \in \{1, 2, 3, 4\}$ will be called proper images obtained by $\sigma_{(\bar{p}, \kappa)}$.

In \mathbb{QH}_2 , by choosing different types of the line-conics for the fundamental line-conic and different positions of the polar line, we can have different types of line-inversion. Because of this variety, and based on the analogue of Euclidean inversion, in the following, two types of line-inversion will be studied in accordance to the type of the fundamental line-conic. Thus we have *circle line-inversion* and *equiform line-inversion*, where based on the type or position

of the polar line we obtain subtypes of certain line-inversion. Furthermore, we will observe the images obtained by the line-inversion that result with a proper line-conic i.e. images of pencils of lines not containing the fundamental lines and line-conics with two lines coinciding with the fundamental lines.

3.1 Circle line-inversion

Let the fundamental line-conic κ be given as a circle whose equation from (12) is of the form

$$\kappa : -u_0^2 + u_1^2 - u_2^2 = 0,$$

with the central line $\bar{c} \sim [1 : 0 : 0]$. Two subtypes of the circle line-inversion will be observed.

Type 1) Circle line-inversion $\sigma_{(\bar{c}, \kappa)}$ with respect to the central line.

This type of line-inversion is an analogue of the Euclidean inversion. The fundamental lines are $\bar{c}, \bar{f}_1, \bar{f}_2$ and the equation of the transformation is given with

$$\sigma_{(\bar{c}, \kappa)} : \bar{u} \rightarrow \bar{v} \sim [u_1^2 - u_2^2 : u_0 u_1 : u_0 u_2].$$

The proper image of the pencil $\mathbf{q}\bar{u} = 0$, where $q_0(q_1^2 - q_2^2) \neq 0$, is the circle with equation

$$q_0 u_1^2 - q_0 u_2^2 + q_1 u_0 u_1 + q_2 u_0 u_2 = 0,$$

having the central line $[-2q_0 : q_1 : -q_2]$.

Line-conics containing two fundamental lines are either circles not containing the polar line, or special hyperbolas and special parabolas containing the polar line. The proper image of the circle given with (13) has the equation

$$a^2 u_0^2 + 4a u_0 u_1 + 2u_1^2 - 2u_2^2 = 0,$$

with the central line $[-1 : a : 0]$. Hence, the generating circle and the obtained circle are concentric. Specially, the circle concentric with κ is given with (12) and its proper image is

$$u_0^2 - a u_1^2 + a u_2^2 = 0,$$

thus they are all concentric (see fig. 4).

Remark 1. The proper image of a circle concentric with κ is also the polar-reciprocal line-conic to the generating curve obtained by π_κ .

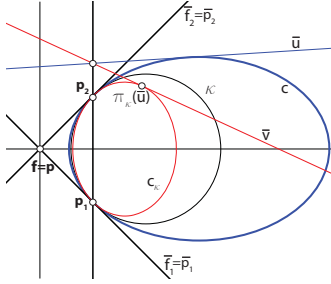


Figure 4: Image c_K of the circle c with respect to $\sigma_{(\bar{c}, \kappa)}$

For a special hyperbola and a special parabola containing the line \bar{p} it follows that $c_{00} = 0$ and the obtained proper image with respect to $\sigma_{(\bar{c}, \kappa)}$ will generally result with either a special hyperbola or special parabola. For example, from (9) the special parabola has the equation $u_1^2 + u_2^2 + 2au_0u_1 + 2bu_0u_2 \pm 2u_1u_2 = 0$, $a, b \in \mathbb{R}$. The proper image is the special hyperbola

$$2au_1^2 \mp 2bu_2^2 + u_0u_1 \pm u_0u_2 + 2(\mp a + b)u_1u_2 = 0,$$

with the central line $[-2(a \pm b)^2 : -a \pm 3b : \mp 3a + b]$ (see fig. 5). If $a = \pm b$ then the image is the special parabola

$$u_0u_1 - u_0u_2 + 2a(u_1 \pm u_2)^2 = 0.$$

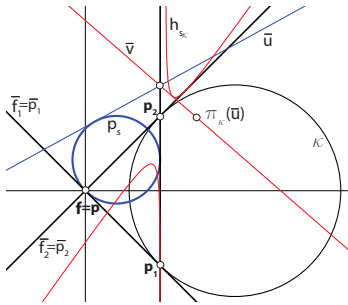


Figure 5: Image h_{s_K} of the special parabola p_s with respect to $\sigma_{(\bar{c}, \kappa)}$

These results are summarized in the following theorem:

Theorem 2. A circle line-inversion $\sigma_{(\bar{c}, \kappa)}$ in \mathbb{QH}_2 maps

- i) pencils of lines not containing the fundamental lines and circles not containing the polar line into circles,
- ii) special hyperbolas and special parabolas containing the polar line into line-conics, which are either special hyperbolas or special parabolas.

Type 2) Circle line-inversion $\sigma_{(\bar{p}, \kappa)}$ with respect to an isotropic line $\bar{p} \sim [0 : p_1 : p_2]$. The isotropic lines \bar{p} and $\bar{r} \sim [0 : p_2 : p_1]$, where $H(\bar{f}_1 \bar{f}_2, \bar{p} \bar{r}) = -1$ are the singular lines of $\sigma_{(\bar{p}, \kappa)}$. In general the equation of the transformation is given with

$$\begin{aligned} \sigma_{(\bar{p}, \kappa)} : \bar{u} \sim [u_0 : u_1 : u_2] &\rightarrow \bar{v} \sim [v_0 : v_1 : v_2] \\ v_0 &\sim -p_1u_0u_1 + p_2u_0u_2, \\ v_1 &\sim -p_1(u_0^2 + u_2^2) + p_2u_1u_2, \\ v_2 &\sim p_2(-u_0^2 + u_1^2) - p_1u_1u_2. \end{aligned}$$

The proper image of the pencil $\mathbf{q}\bar{u} = 0$ is the line-conic of form

$$\begin{aligned} &p_2(u_0u_2q_0 - u_0^2q_2 + u_1(u_2q_1 + u_1q_2)) - \\ &p_1(u_0u_1q_0 + u_0^2q_1 + u_2(u_2q_1 + u_1q_2)) = 0. \end{aligned}$$

Two subtypes will be observed, whether the polar line is an arbitrary isotropic line or one of the absolute lines.

If we choose $\bar{p} \sim [0 : 0 : 1]$, then the other two fundamental lines are non-isotropic lines $\bar{p}_1 \sim [1 : 1 : 0]$, $\bar{p}_2 \sim [1 : -1 : 0]$. The equation of the transformation is

$$\sigma_{(\bar{p}, \kappa)} : \bar{u} \rightarrow \bar{v} \sim [u_0u_2 : u_1u_2 : -u_0^2 + u_1^2].$$

The proper image of the pencil $\mathbf{q}\bar{u} = 0$, where $q_2(q_0^2 - q_1^2) \neq 0$ is in general the hyperbola with the equation

$$u_0u_2q_0 + u_1u_2q_1 + q_2(-u_0^2 + u_1^2) = 0,$$

having the central line $[-q_1^2 : q_0q_1 : -2q_0q_2]$. There are two exceptions, the first is when \mathbf{q} is an isotropic point ($|q_1| = |q_2| = 1$) then the obtained line-conic is the special hyperbola with the equation

$$\pm(-u_0^2 + u_1^2) + q_0u_0u_2 + \pm u_1u_2 = 0,$$

having the central line $[1 : \pm q_0 : 2q_0]$.

The other exception occurs when $\mathbf{q} \sim (1 : 0 : q_2)$ is incident with the line \bar{r} , then the obtained line-conic is the parabola with the equation

$$u_0u_2 - u_0^2q_2 + u_1^2q_2 = 0,$$

having the central line $[0 : 0 : 1]$.

If the generating curve is a line-conic containing two fundamental lines two cases occur, either it contains \bar{p} , \bar{p}_1 or \bar{p}_1 , \bar{p}_2 . In the first case, the coefficients of the line-conic (2) satisfy $c_{22} = 0$, $c_{00} + c_{11} + 2c_{01} = 0$ which in general is a hyperbola. For $c_{11} \pm 2c_{12} = 0$ it is the special hyperbola or for $c_{12} = 0$ the parabola.

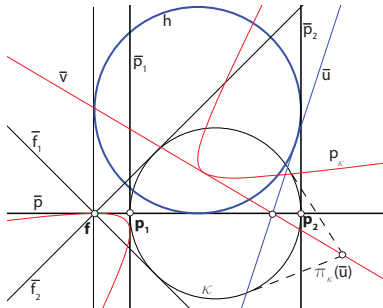


Figure 6: Image p_K of the equiform hyperbola h with respect to $\sigma_{(\bar{p}, \kappa)}$

Generally, the proper image is given with the equation

$$2(u_0 + u_1)(c_{02}u_0 + c_{12}u_1) + (2c_{01}u_0 + c_{11}(u_0 + u_1))u_2 = 0, \quad (24)$$

which is either the hyperbola, special hyperbola or parabola. From the conditions for the coeffi-

cients in (24) it can be concluded that only special hyperbolas can be obtained from special hyperbolas, and that the parabolas are obtained if $c_{11} = 0$, hence when the generating line-conics are equiform hyperbolas (see fig. 6).

In the second case the coefficients of the line-conic satisfy $c_{00} = -c_{11}$, $c_{01} = 0$ which can be any line-conic and its proper image is

$$c_{22}(u_0 - u_1)(u_0 + u_1) + u_2(-2c_{02}u_0 - 2c_{12}u_1 + c_{00}u_2) = 0, \quad (25)$$

with the central line $[c_{12}^2 + c_{00}c_{22} : -c_{02}c_{12} : c_{02}c_{22}]$. From the conditions for coefficients in (25) it can be concluded that the obtained proper line-conics are of the same type as their generating line-conics (see fig. 7).

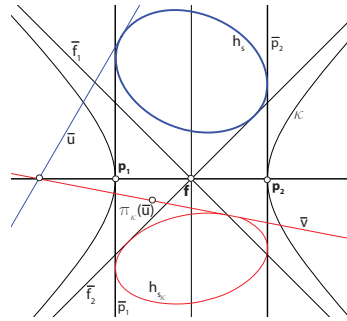


Figure 7: Image h_{s_K} of the special hyperbola h_s with respect to $\sigma_{(\bar{p}, \kappa)}$

If the polar line is the absolute line (for instance $\bar{p} = \bar{f}_1$) then all fundamental lines coincide with that absolute line. The equation of the transformation is of the form

$$\begin{aligned} \sigma_{(\bar{f}_1, \kappa)} : \bar{u} \sim [u_0 : u_1 : u_2] &\rightarrow \bar{v} \sim [v_0 : v_1 : v_2] \\ v_0 &\sim u_0(-u_1 + u_2), \\ v_1 &\sim -u_0^2 + (u_1 - u_2)u_2, \\ v_2 &\sim -u_0^2 + u_1(u_1 - u_2). \end{aligned}$$

The proper image of the pencil $\mathbf{q}\bar{u} = 0$, where $q_1^2 - q_2^2 \neq 0$, is the special hyperbola with the

equation

$$u_0(-u_1 + u_2)q_0 - u_0^2(q_1 + q_2) + (u_1 - u_2)(u_2q_1 + u_1q_2) = 0,$$

having the central line $[(q_0 - 4q_1)(q_2 - q_1) : q_0(-3q_1 + q_2) : -q_0(q_0 + 2q_1 - 2q_2)]$. If \mathbf{q} is an isotropic point ($q_1 = q_2 = \pm 1$) the proper image is the circle with the equation

$$\pm 2u_0^2 \mp u_1^2 + \pm u_2^2 + u_0(-u_1 + u_2)q_0 = 0$$

with the central line $[\mp 2 : q_0 : q_0]$ (see fig. 8).

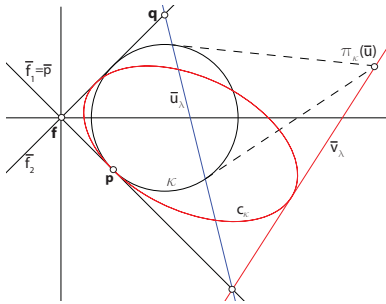


Figure 8: Image c_κ of the pencil (\mathbf{q}) with respect to $\sigma_{(\bar{f}_1, \kappa)}$

If the generating curve is a line-conic containing the polar line as a double line i.e. the absolute line \bar{f}_1 as a 2-fold line, then the coefficients of the line-conic (2) satisfy $c_{01} = -c_{02}$ and for $a, b, c \in \mathbb{R}$ the equation is

$$u_0^2 + au_1^2 + bu_2^2 + 2c(u_0u_1 - u_0u_2) - (a+b)u_1u_2 = 0,$$

which is either a special hyperbola containing \bar{f}_1 or a circle for $a = -b$. The obtained proper line-conic has the equation

$$(1 + a - b)u_0^2 + 2cu_0(u_1 - u_2) + (u_1 - u_2)(bu_1 - au_2) = 0,$$

which is again either a special hyperbola containing the line \bar{f}_1 or a circle for $a = -b$.

Theorem 3. A circle line-inversion $\sigma_{(\bar{p}, \kappa)}$ in \mathbb{QH}_2 , where the polar line \bar{p} is an isotropic line, maps

i) pencils of lines not containing the fundamental lines into either hyperbolas, special hyperbolas and parabolas or into special hyperbolas and circles if \bar{p} is an absolute line,

ii) line-conics containing two fundamental lines into either hyperbolas, special hyperbolas and parabolas if the generating line-conics contain the polar line or into line-conics of same type as the generating line-conic if it does not contain the polar line. If \bar{p} is an absolute line then special hyperbolas and circles are mapped into special hyperbolas and circles respectively.

3.2 Equiform line-inversion

Let the fundamental line-conic be given either as the equiform hyperbola with the equation from (16)

$$\kappa_H : -u_0^2 + 2u_1u_2 = 0$$

or as the equiform ellipse with the equation from (17)

$$\kappa_E : -u_0^2 + u_1^2 + u_2^2 = 0.$$

For both, the central line is $\bar{c} \sim [1 : 0 : 0]$.

We will observe the equiform line-inversions $\sigma_{(\bar{c}, \kappa_H)}$, $\sigma_{(\bar{c}, \kappa_E)}$ when the polar line is the central line. The equations of the transformations are of the form

$$\sigma_{(\bar{c}, \kappa_H)} : u \rightarrow v \sim [2u_1u_2 : u_0u_1 : u_0u_2]$$

$$\sigma_{(\bar{c}, \kappa_E)} : u \rightarrow v \sim [u_1^2 + u_2^2 : u_0u_1 : u_0u_2].$$

This type of line-inversion maps the absolute lines onto each other.

In both cases the proper image of the pencil $\mathbf{q}\bar{u} = 0$ is a equiform line-conic. In the first case it is the equiform hyperbola with the equation

$$2u_1u_2q_0 + u_0u_1q_1 + u_0u_2q_2 = 0$$

having the central line $[-2q_0 : q_2 : q_1]$ (see fig. 9) and in the second case the equiform ellipse with the equation

$$(u_1^2 + u_2^2)q_0 + u_0u_1q_1 + u_0u_2q_2 = 0$$

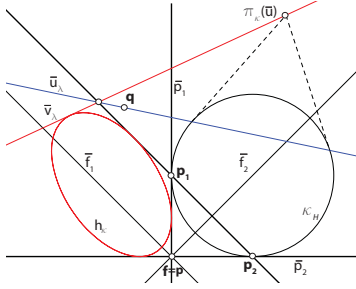


Figure 9: Image h_K of the equiform hyperbola h with respect to $\sigma_{(\bar{c}, \kappa_H)}$

having the central line $[2q_0 : -q_1 : -q_2]$.

For $\sigma_{(\bar{c}, \kappa_H)}$ the other two fundamental lines are $\bar{p}_1 \sim [0 : 0 : 1]$ and $\bar{p}_2 \sim [0 : 1 : 0]$. Hence, there are two cases of a line-conic containing two fundamental lines, either it contains \bar{c} , \bar{p}_1 or \bar{p}_1 , \bar{p}_2 . In the first case the coefficients of the line-conic satisfy $c_{00} = c_{22} = 0$, which in general is a hyperbola. For $c_{11} = \pm 2c_{12}$ it is the special hyperbola and for $c_{12} = 0$ the parabola. The image is of the form

$$c_{11}u_0u_1 + 2c_{12}u_0u_2 + 4c_{01}u_1u_2 + 4c_{02}u_2^2 = 0,$$

with the central line $[2c_{01}^2c_{11} : c_{01}c_{11}c_{12} - 2c_{02} : c_{01}]$ which, in general, is the hyperbola. For $c_{01} = \pm c_{02}$ it is the special hyperbola and for $c_{01} = c_{02} = 0$ it is the parabola.

In the second case the coefficients of the line-conic satisfy $c_{11} = c_{22} = 0$, hence they are equiform hyperbolas. The proper image of the equiform hyperbola given with (18) is the equiform hyperbola with the equation

$$a^2u_0^2 + 2au_0u_1 + 2au_0u_2 + 2u_1u_2 = 0,$$

having the central line $[-2 : a : a]$. Specially, the equiform hyperbola given with (16) is concentric with κ_H and its image is the equiform hyperbola with the equation $u_0^2u_1u_2 + 2au_1^2u_2^2 = 0$ also concentric with κ_H .

For $\sigma_{(\bar{c}, \kappa_E)}$ the other two fundamental lines are conjugate complex and therefore only the

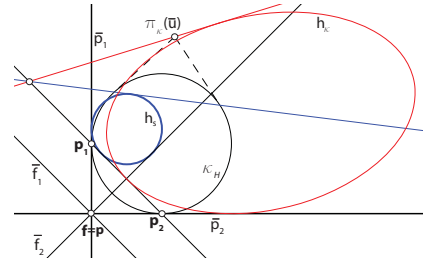


Figure 10: Image h_K of the special hyperbola h_s with respect to $\sigma_{(\bar{c}, \kappa_H)}$

equiform ellipses (17) concentric with the fundamental ellipse κ_E obtain a proper real line-conic. Its image is a equiform ellipse with the equation

$$-u_0^2 + au_1^2 + au_2^2 = 0$$

and it is also concentric with κ_E .

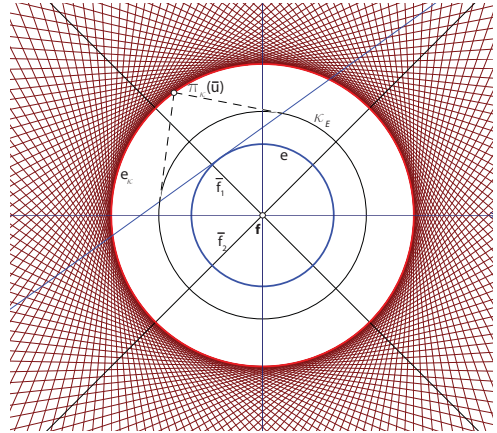


Figure 11: Image e_K of the ellipse e with respect to $\sigma_{(\bar{c}, \kappa_E)}$

Remark 2. The proper image of an equiform line-conic concentric with κ_H or κ_E is also the polar-reciprocal line-conic to the generating curve obtained by π_{κ_H} or π_{κ_E} respectively.

Theorem 4. An equiform line-inversion $\sigma_{(\bar{c}, \kappa_H)}$,

$\sigma_{(\bar{c}, \kappa_E)}$ in \mathbb{QH}_2 , where κ_H is the equiform hyperbola and κ_E is the equiform ellipse, maps

- i) pencils of lines not containing the fundamental lines and equiform line-conics central with the fundamental line-conic into the equiform line-conics of the same class as the fundamental line-conic,
- ii) line-conics containing the polar line and one other fundamental line with respect to $\sigma_{(\bar{c}, \kappa_H)}$ into the hyperbolas, special hyperbolas or parabolas.

4. CONCLUSIONS

This paper presents two different types of the line-inversion in the quasi-hyperbolic plane i.e circle line-inversion and equiform line-inversion. The polar line for the circle line-inversion is chosen as the central line of the fundamental circle or an arbitrary isotropic line. The first choice is an analogue of the well known Euclidean inversion, and the second choice is a special case of the generalised inversion in Euclidean plane. The focus of this paper is onto the description of the proper line-conics obtained by these types of the line-inversion.

REFERENCES

- [1] T. A. Hirst. On the Quadric Inversion of Plane Curves. *Proceedings of the Royal Society of London*, 14: 91–106, 1865.
- [2] N. Kovačević, V. Szivoczka. Inversion in Minkowskischer geometrie. *Mathematica Pannonica*, 21/1: 89–113, 2010.
- [3] V. Niče. Curves and surfaces of 3rd and 4th order deduced by quadratic inversion. *Rad HAZU*, 278: 153–194, 1945. (in Croatian)
- [4] H. Sachs. *Ebene Isotrope Geometrie*. Friedr. Vieweg & Sohn, Braunschweig/Wiesbaden, 1987.
- [5] A. Sliepčević, I. Božić, H. Halas. Introduction to the planimetry of the quasi-hyperbolic plane. *KoG*, 17: 58–64, 2013.
- [6] I.M. Yaglom. *Galilean principle of relativity and NonEuclidean geometry*. Nauka, Moscow, 1969.
- [7] I.M. Yaglom, B.A. Rozenfeld, E.U. Yasin-skaya. Projective metrics. *Russ. Math Surveys*, 19/5: 51–113, 1964.

ABOUT THE AUTHORS

1. Helena Halas, M. Sc., is Research Assistant at the Faculty of Civil Engineering (Department of Mathematics) at the University of Zagreb. Her research interests are projective geometry, Euclidean and non-Euclidean planes, treated by synthetic (constructive) and analytical methods. She can be reached by e-mail: hhalas@grad.hr.
2. Nikolina Kovačević, M. Sc., is Lecturer at the Faculty of Mining, Geology and Petroleum Engineering (Department of Mathematics, Informatics and Descriptive Geometry) at the University of Zagreb. Her research interest are projective geometry, Euclidean and non-Euclidean geometries, and descriptive geometry. She can be reached by e-mail: nikolina.kovacevic@rgn.hr
3. Ana Sliepčević, Ph. D., is retired Associate Professor. Her research interest are projective geometry, Euclidean and non-Euclidean planes treated by synthetic methods. She can be reached by e-mail: anas@grad.hr

THE LOCUS-FUNCTION OF THE PERSPECTIVE CIRCLE

Michael SEJFRIED

METAL UNION, Czestochowa, Poland

ABSTRACT: This paper presents the locus-curve of the central points of the base-circle. We have two main points and we look for the position (connected to these points) of the special circle, which should fulfill certain conditions. These conditions are determined by six points lying on the circle and nine lines connecting these points. So we get three triples of the lines intersecting at three points. These intersections should be collinear. Seeking similar geometrical problems I didn't find any materials which would be able to help me in this elaboration, I came across no information referring to similar constructions. That is why references to this paper contain only two positions which I mainly used. The locus-curve, the circle and the mentioned lines have a few very interesting properties. I will try to describe them in this paper.

Keywords: locus, curve, circle, triangles, equality.

1. INTRODUCTION

This paper presents the locus-curve of the central points of the base-circle. We have two main points and we look for the position (connected to these points) of the special circle, which should fulfill certain conditions. These conditions are determined by six points lying on the circle and nine lines connecting these points. So we get three triples of the lines intersecting at three points. These intersections should be collinear. Seeking similar geometrical problems I didn't find any materials which would be able to help me in this elaboration, I came across no information referring to similar constructions. That is why references to this article contain only two positions which I mainly used. The locus-curve, the circle and the mentioned lines have a few very interesting properties. I will try to describe them in this paper. I will try to introduce the Sejfriedian study in training courses on geometry.

2. CONSTRUCTION

Given is reference line segment **AB** and its length **d**. For simplicity's sake let's lie the segment **AB** on the X-axis and the point **A** to be the coordinate origin. Now, the circle $o(O; r)$, which will be called base-circle or perspective circle, lies over the segment **AB**. Let's choose the point **L** on the "right side" of this circle and connect it with the points **A** and **B** (Figure 1).

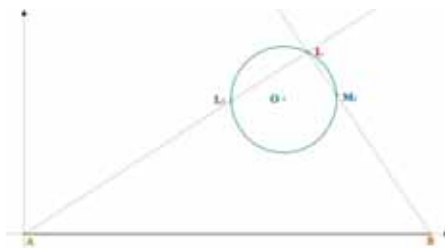


Figure 1: Base-circle and segment.

The line **AL** intersects the circle at the second point **L₁** and the line **BL** – at the second point **M₁**. Let's connect the point **L₁** with the point **B** and the point **M₁** with the point **A**. We get next two intersections with the circle – the points **M** and **K** (Figure 2).

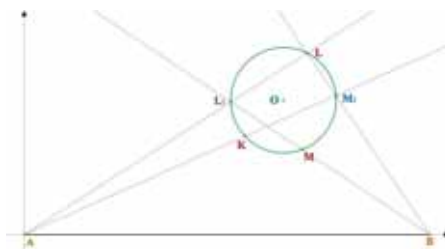


Figure 2: Connection of the points.

The last intersection **K₁** lies on the line **BK** (Figure 3).

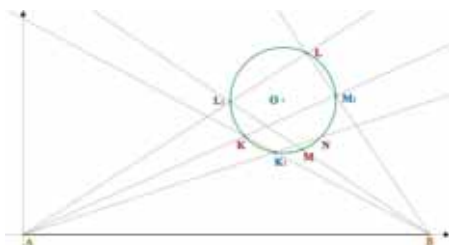


Figure 3: Last intersection.

Generally the point K_1 doesn't lie on the line AM . We must move the center O of $o(O; r)$ vertically so, that we will get above coincidence.

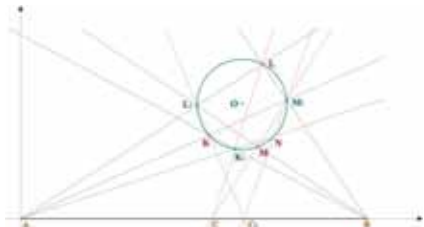


Figure 4: We get coincidence.

Then the points $N \rightarrow M$ and $C \rightarrow C_1$ (Figure 4). The point O determines the locus-curve we are looking for (ζ_0 in the Figure 5).



Figure 5: Locus curve.

The function ζ_0 shown above is determined in the interval:

$$x_0 \in \left(x_A - |AB| \cdot \frac{2\sqrt{3} - 3}{6}, x_B + |AB| \cdot \frac{2\sqrt{3} - 3}{6} \right)$$

and takes there the values:

$$y_0 \in \left(r, \sqrt{\frac{|AB|^2}{12} + r^2} \right).$$

The locus ζ_0 and the whole construction have many interesting properties. Denoting the distances of the points K, L, M, K_1, L_1 and M_1 to the line AB as $h_K, h_L, h_M, h_{K_1}, h_{L_1},$ and h_{M_1} , we get:

$$\begin{aligned} & \frac{1}{|KL|^2} + \frac{1}{|LM|^2} + \frac{1}{|MK|^2} = \\ & = \frac{1}{|K_1L_1|^2} + \frac{1}{|L_1M_1|^2} + \frac{1}{|M_1K_1|^2}, \end{aligned}$$

$$\begin{aligned} |KK_1| \cdot |L_1M| \cdot |LM_1| &= |LL_1| \cdot |M_1K| \cdot |MK_1| = \\ &= |MM_1| \cdot |K_1L| \cdot |KL_1|, \end{aligned}$$

$$\alpha_L + \beta_L = \gamma_R, \text{ and } \alpha_R + \beta_R = \gamma_L.$$

After-mentioned equations hold for given x_0 independent on the position of the point L . The point L can move between the points A_3 and B_2 – Figure 5. During this movement the values of $h_K, h_L, h_M, h_{K_1}, h_{L_1}$, and h_{M_1} , change, but nonetheless we get following equations:

$$\begin{aligned} \frac{1}{h_K} + \frac{1}{h_L} + \frac{1}{h_M} &= \frac{1}{h_{K_1}} + \frac{1}{h_{L_1}} + \frac{1}{h_{M_1}} = v_1 \\ &= \text{const}, \frac{1}{h_K^2} + \frac{1}{h_L^2} + \frac{1}{h_M^2} \\ &= \frac{1}{h_{K_1}^2} + \frac{1}{h_{L_1}^2} + \frac{1}{h_{M_1}^2} = v_2 \\ &= \text{const}. \end{aligned}$$

Denoting the distances of the points A_2, A_3, B_2, B_3, C_2 and C_3 to the line AB as $h_{A_2}, h_{A_3}, h_{B_2}, h_{B_3}, h_{C_2}$ and h_{C_3} we get: There is also:

$$\begin{aligned} \frac{1}{h_{A_2}} + \frac{1}{h_{B_2}} + \frac{1}{h_{C_2}} &= \frac{1}{h_{A_3}} + \frac{1}{h_{B_3}} + \frac{1}{h_{C_3}} = v_1 \\ &= \text{const}, \\ \frac{1}{h_{A_2}^2} + \frac{1}{h_{B_2}^2} + \frac{1}{h_{C_2}^2} &= \frac{1}{h_{A_3}^2} + \frac{1}{h_{B_3}^2} + \frac{1}{h_{C_3}^2} = v_2 \\ &= \text{const}. \end{aligned}$$

The function ζ_0 and its equation I found for about 5 years. The equation I got that time using the program *Mathematica* was long and complicated:

$$\begin{aligned} l &= |AB|, x = x_0, v \\ &= \sqrt{l^2 + 12 \cdot l \cdot x - 12x^2}, v_1 \\ &= r^2 - x^2, v_2 = r^2 + 3x^2, v_3 \\ &= 6lxv_1 + 3v_1^2 + l^2v_2. \\ u_1 &= l^2 + 18r^2 + 6x(l - x) - lv + \frac{3r^4}{v_2} \\ &+ \frac{3lr^4v}{v_3}, u_2 \\ &= u_1 + \frac{9r^4(r^4 + 6r^2x^2 + x^3(8l - 7x))}{v_2v_3}, \\ u_3 &= 18x^4v_1 + l^4v_2, \\ u_4 &= u_3 + l^3(3x^2(4x - v) + r^2(12x - v)), \end{aligned}$$

$$\begin{aligned}
u_5 &= u_4 - 3lx^2(2r^2(6x - v)) - x^2(18x \\
&\quad - v), u_6 \\
&= u_5 + 3l^2x(2r^2(x - v) \\
&\quad - x^2(17x - 2v)), \\
y_0 &= \frac{1}{2\sqrt{6}} \left(\sqrt{u_2} + \sqrt{\frac{u_6}{v_3}} \right).
\end{aligned}$$

In March 2014 Vladimir Shelomovskii paid attention that this function has still the second branch lying above the given by this equation. Changing the sign of v we get:

$$v = -\sqrt{l^2 + 12 \cdot l \cdot x - 12x^2},$$

and the upper branch of the function ζ_0 . Vladimir Shelomovskii found later the following parametric equation for this function:

$$x = \frac{L}{2} + \frac{L}{\sqrt{3}} \cos \phi,$$

$$y^2 = r^2 + \frac{L^2}{12} (1 + 2 \sin \phi)^2,$$

where $L = |\mathbf{AB}|$. Let's convert it to the function $\zeta_0(x)$:

$$\begin{aligned}
&\zeta_0(x) \\
&= \sqrt{r^2 + \frac{L^2}{12} \left(1 \mp 2 \sqrt{1 - \frac{3(2x - L)^2}{4L^2}} \right)^2}.
\end{aligned}$$

The function ζ_0 has two branches: blue for the sign $(-)$ and red for the sign $(+)$ in above equation. In below figures are shown 3 examples: $L = 20$ and $r = \{3; 1; 0\}$.

For $r = 0$ we get the function:

$$\zeta_0(x) = \frac{1}{2\sqrt{3}} |L \mp \sqrt{L^2 + 12Lx - 12x^2}|.$$

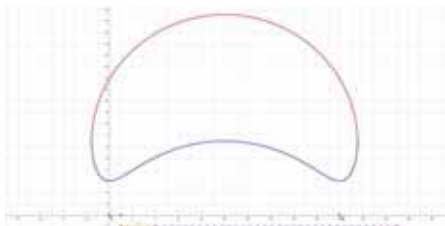


Figure 6: Curve for $r = 3$.



Figure 7: Curve for $r = 1$.



Figure 8: Curve for $r = 0$.

which is the circle with the part under the line \mathbf{AB} reflected in respect to this line \mathbf{AB} .

The base-circle moving along the curve $\zeta_0(x)$ fulfills all the conditions given at the begin. Changes only the order of the points K , L , M , K_1 , L_1 and M_1 on the circle.

3. CONCLUSIONS

Interesting is the considering, whether similar task could have the solutions in respect to three not collinear points A , B and C lying at the coordinates origin of the XYZ -system and on the axes X and Y . The point C from 2D would be replaced by the point D and the base-circle – by the base-sphere.

REFERENCES

- [1] E.W. Weisstein, CRC Concise Encyclopedia of Mathematics, Chapman & Hall, 2003.
- [2] www.mathworld.wolfram.com.

ABOUT THE AUTHOR

Michael Sejfried is not professional mathematician. He is dealing with geometry in his spare time. Affiliation: the company METAL UNION in Poland. He can be reached by email via the address: michael@sejfried.pl.

LOGIC (SYSTEMATIC) VS. AUTOMATION (MECHANISM)

Antonio MOLLICONE
University of Malta, Malta

ABSTRACT: The teaching of descriptive geometry and its applications is fundamental in courses for designers that will have to foresee, in the field of the physical space conceived in three linear dimensions, objects related to each other.

They will have to express themselves through forms/shapes and in general with the construction of solids in whichever way these may be placed and oriented.

The construction of graphical models of objects conceived in a design phase share, through projections operations, both the choice of position in the reference system and the consideration of non-finite geometrical elements introduced by projective geometry.

The software substitutes the drawing tools with pre-disposed commands that enable to visualise, in a virtual space, all the graphical operations.

The use of the software tool is, inevitably, a shortcut to solve problems in a fast and accurate way even thanks to the capability of simulating space.

Unfortunately this prerogative has greatly diminished the practice of traditional drawing and has convinced many on its backwardness, so that what is witnessed is an empirical, approximative, use of the software tool based on the knowledge of automatic commands.

Problems in descriptive geometry are not examined through formulae, but understanding the spatial relationship of the geometrical entities subject of the study at hand and hence through the choice of the logical route that, in the simplest way, will lead to the solution.

In descriptive geometry the solution of the most notable problems is not only a theoretical exercise.

Knowing how to control the rotation of a solid, defining the apparent contour implies an intimate knowledge of it so as to use it adequately in the design phase.

The need to unify programs (ECTS) together with the wide-spread idea of the computer as the solution for all representation problems lead to compress the geometrical-descriptive teaching in the hotchpotch of the “Design Workshop” confining it to the role of graphical support more and more aimed at the pure photo-realistic effect.

Architectural design is also a geometrical exercise and the relative representation an application of descriptive geometry that, through the solution of problems, better if located in construction details, prepares students to the professional practice like the inverse route of surveying exercises.

The various methods of representation should be illustrated in parallel considering them similar for the common elements of the projection and every demonstration illustrated in all the logical steps: students, in general, come from a high school whose syllabus does not always include the teaching of plane and solid geometry. This contribution illustrates the graphical solution of problems that, even in the software model, find a solution only through a logical process leaving to the accuracy of the software tool the verification of the result. (Fig. 1)

Keywords: Descriptive geometry, projective geometry, planar and solid geometry.

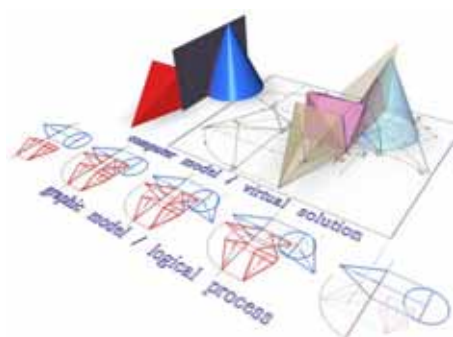


Figure 1: Graphic and computer models.

1. THE SCIENTIFIC ILLUSTRATION

It turns out to be difficult to recognize, in the drawing of Fig. 2 [1], the solution of the problem to represent, in central projection, the perpendicular brought from a point A to line b both belonging to a generic plane.

In the virtual environment (Fig. 3) the geometrical elements of the problem can be visualized and with an automatic view a more understandable image (Fig.4) is produced and adapted for students of architectural schools.

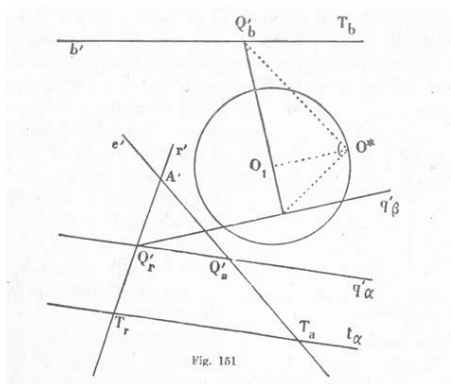


Figure 2: Perpendicular brought from a point A to line b .

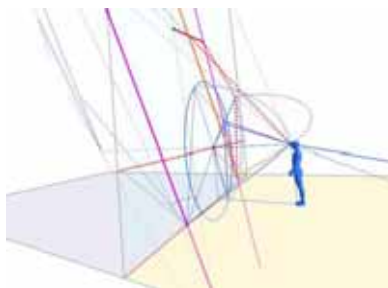


Figure 3: Visualization of the geometric elements of a problem.

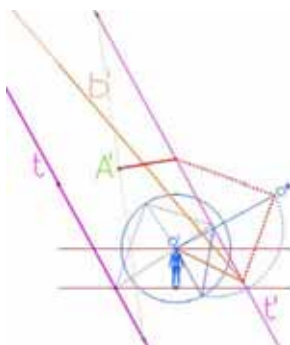


Figure 4: Orthographic projection from 3d model.

2.THE TEACHING OF DESCRIPTIVE GEOMETRY

The applications of elementary geometry of space and of projective geometry, fundamentals of descriptive geometry, have been taught with more realistic illustrations (Fig. 5) [2], and after the contribution of Prof. O. Fasolo (Fig. 6) [3], they have been proposed through the computer tool in the school of Prof. R. Migliari (Fig.7) [4].

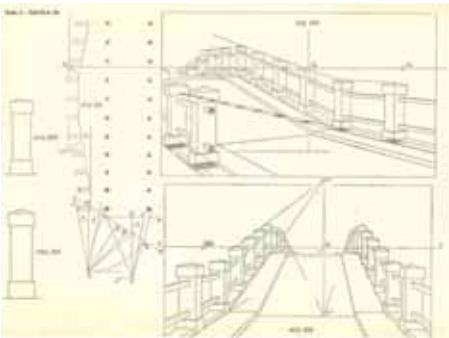


Figure 5: Measurement in perspective: indirect method.

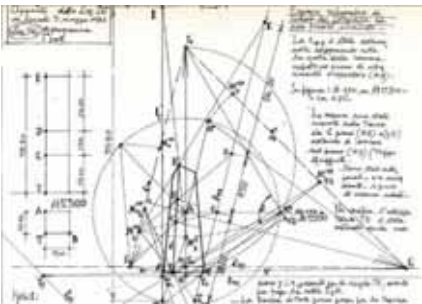


Figure 6: Measurement in perspective: direct method.

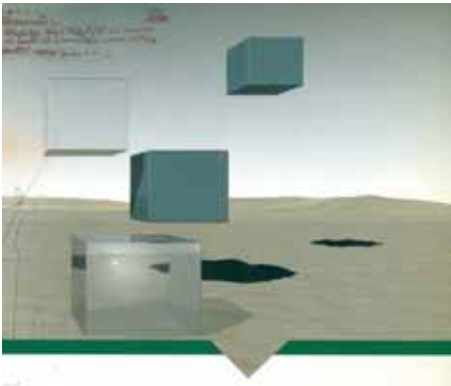


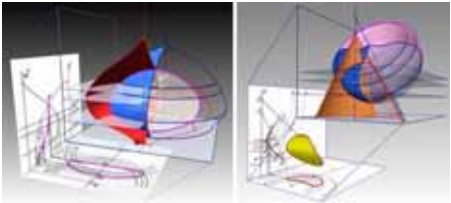
Figure 7: Computer model: perspective view.

3. THE PRACTICALITY OF ORTHOGRAPHIC PROJECTIONS

“The representation of spatial figures is usually obtained through projection and sections;” and with the use of planes and auxiliary figures opportunely placed and rotated some problems apparently complex are solved in the graphical model with the appropriate tools of the Architect: the ruler and the compass.

The computer tool definitively “solves” the problem of representation and hence assures an absolute divulgation and comprehension.

(Fig. 8-9). The infinite views and animations provided by the software are to be compared with the complete representation provided by the simultaneous views in the method of the triple orthogonal projections. (Fig. 10)



Figures 8-9 :Intersection of rotational surfaces.

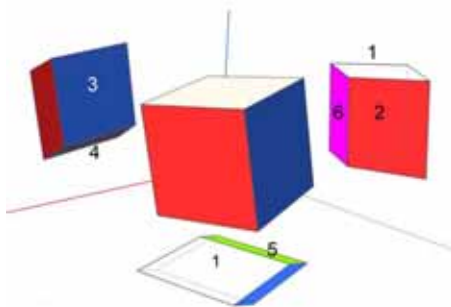


Figure 10: Full view of the object in triple orthogonal projections.

4. THE LINEARITY OF ORTHOGRAPHIC PROJECTIONS

In the graphical model the solution of problems requires a relatively short time if the relationship with relevant elementary geometrical solutions is perceived and opportune views and series of operations anyway easily representable are chosen.

In the computer model we go through, in the virtual space, the logical process and we verify the exactness/accuracy of the solution with a superior illustrative efficiency. (Fig. 11)

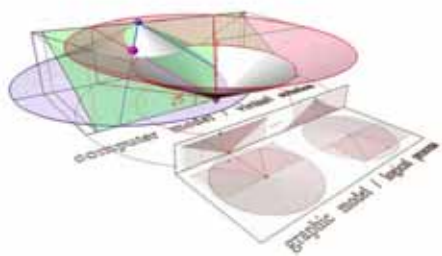
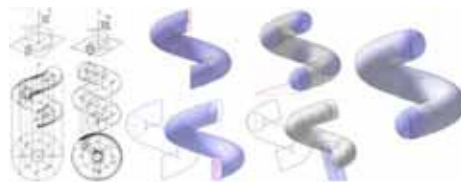


Figure 11: Comparison between graphic and computer model.

5. THE COMPUTER MODEL AS SUPPORT FOR REPRESENTATION

The problem of the representation of certain surfaces is solved with the enveloping of the projections of sections of the same one but only in the computer model the “limit points” of the internal branch of the apparent contour can be visualized (light contours) and defined (projections on a plane) [5].

As an example, imagine the surface of the torus transformed in that of rotation-translation like the “screw of Saint Gilles” and “serpentine”. (Fig. 12 - 13)



Figures 12-13: Graphic and computer models: rotation-translation surfaces.

6. THE VERIFICATION THROUGH THE COMPUTER MODEL

The curve of apparent contour of a surface of rotation-translation cannot coincide completely with the meridian section of the same one and in the case of the “twisted column” it is difficult to “imagine it” and represent it graphically. With the computer model it can be determined automatically and it is verified that the mentioned apparent contour is not the meridian curve but the helix on the cylinder that has as diameter the dimension of the projection of the meridian curve i.e. the difference between the diameter of the circle that envelops the projections of the generatrix and the diameter of the same generatrix. (Fig. 14)

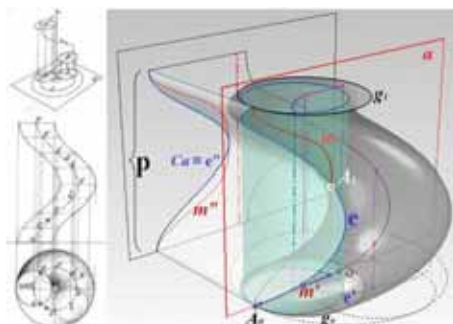


Figure 14: Solution and verification through the computer model.

7. PHOTOREALISM AND ANIMATION INCLUDING DEDUCTION AND LOGIC

The computer techniques of chiaro-scuro, texture and animation enable us to build models that, in the virtual world, can represent effectively reality and hence verify the consequences of hypothetical interventions on the same reality. Without doubt these tools are important because they solve definitively the problem in the field of communication / spreading of ideas. In the teaching/didactic field they offer the possibility to illustrate the phases of the process of solution of spatial problems that the student will face in the design phase. (Fig. 15)

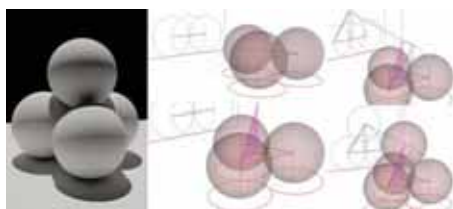


Figure 15: Close-packing of equal spheres. Hexagonal close-packed- hcp.

8. THE EXPERIENCE OF DESIGN FROM THE IMAGINATION TO MEASUREMENT

For the designer, the geometrical operations at the heart of the creative process, have a significant importance because, on one hand he will always have a reference to the origin of the planning, and on the other hand he will take suggestions for solutions that will highlight the main/supporting structure of the idea.

The automatic techniques have reduced the times to realize the elaborate graphics but at the same time can induce to the simplifying practice of producing images loaded with effects.

The ability to “build” graphical models of whatever spatial organization with the practical method of the double and triple orthogonal projections, learnt from the theoretical teaching, is fundamental for the effective control of the design idea.

The aim is not only to appear, rise but build, through drawing, a structure defining the measurements and the reciprocal positions of the elements with respect to the operational field.

The principle that the designer has to use only the compass and the ruler is reinforced, taken as tools but especially as a mental setting.

A suggestion as an aid to the above mentioned principle is that design and verification must make use of the perspective views.

To this end the computer tool proves to be practical and efficient.

The attitude, in my opinion, that has to be fed to the students is that to guess/perceive, looking at an image or reality itself, the part hidden by the apparent contour in order to imagine the elements of one's own design in their complexity. (Fig. 16)

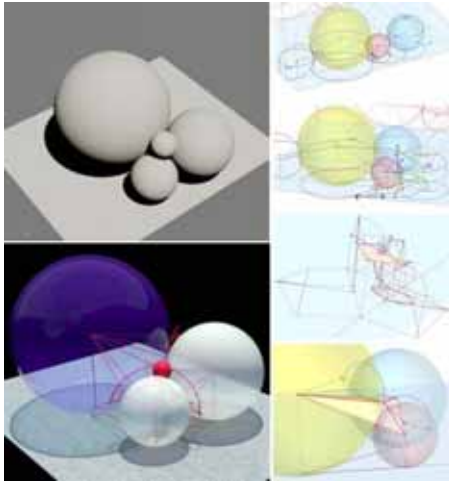


Figure 16: Close-packing of different spheres.

9. CONCLUSIONS

Digital techniques facilitate the visualization of any composition concept and the construction of the relative iconographic and structural models.

In the virtual space architects, being able to “model” the components of the project, regain the experience of the sculptor increasing the plastic values and controlling the visual effect.

REFERENCES

- [1] O.Chisini-G. M. Biggioggero.
Lezioni di Geometria Descrittiva.
Milano. TamburiniEditore.
341 p., Settima edizione 1975.
- [2] F. Vagnetti.
Prospettiva lineare.
Roma - Casa Editrice Mediterranea
II volume - 187 p.. 1949.
- [3] Orseolo Fasolo.
Lezioni di geometria Descrittiva.
Roma- Facoltà di Architettura, 1971
- [4] Riccardo Migliari.
Geometria Descrittiva.
Novara. Città Studi Edizioni.
Volume I - 368 p., 2009.
- [5] Antonio Mollicone .
*La teoria delle ombre lineari e del
chiaro-scuro* - Editrice Kappa
385 p., 2011.

ABOUT THE AUTHOR

1. Antonio Mollicone, born in Rome 1950,

First Degree in Architecture 1973,
Università degli Studi di Roma-Facoltà
di Architettura,

Architect (Warrant) 1974,
Istit. Universitario di Architettura
-Venezia,

PhD Degree (*EUR: Doctor Europaeus*) in
Scienze della Rappresentazione e del
Rilievo-"La Sapienza" University of Rome.

Senior Lecturer at Faculty for the Built
Environment,
Department of Architecture & Urban
Design,
University of Malta,
Coordinator : Course Design I.
antonio.mollicone@um.edu.mt

LOGICAL-CONSTRUCTIVE TEACHING APPROACH THE THEME “CREATION OF SURFACES”

Natalya KAYGORODTSEVA and Tatyana KAYGORODTSEVA

Omsk State Technical University, Russia

ABSTRACT: In this paper authors showed innovative teaching approach to the theme “Creation of Surfaces”. This approach is based on the mathematical calculations of the sum of dimensions of geometrical objects which participate in creation of a surface and dimensions of conditions of their relative positioning. This approach made a mathematical basis for the main theme in discipline “Descriptive geometry”. Earlier this subject was difficult to understand. Students apply logic and therefore logical thinking develops. Future engineers will apply creative approaches to designing surfaces and develops inventive skills which are important in engineering.

Keywords: Surface, Dimension, Descriptive Geometry, Construction.

1. INSTALLATION

All that surrounds us: buildings, machines, household items and others – are geometric solids. Therefore the theme “Creation of surfaces” in Descriptive Geometry is very important. Schoolchildren learn only general meanings about the variety of surfaces at geometry lessons. University students in classical Descriptive Geometry traditionally study solids and surfaces from their representation in the multi-view drawing.

Today, we may create virtual 3D-models of surfaces and study their characteristics with the help of computers and CAD-systems. Therefore we need to change a teaching methods of theme “Creation of surfaces” of Descriptive Geometry. It is necessary to make studying this theme similar to process of creation of surfaces by CAD-systems. Creation of surfaces in Descriptive Geometry needs to be studied taking into account dimensions of conditions of surface determinant. It will enable us to modernize and adapt the course of Descriptive Geometry under the existing realities of computer time.

2. WEAKNESSES OF THE TRADITION-

AL TRAINING COURSE OF DESCRIPTIVE GEOMETRY

The course of Descriptive Geometry traditionally begins with a demonstration of types of projection and the story of Monge's method. Thus, the basic geometrical concept “Point” is not defined. It is supposed that it is known to students from school. But it not so!

Traditional course of Descriptive Geometry does not include a discussion about dimension of a point, the concept “dimension” and the concept “set”. However, a traditional course of Descriptive Geometry describes that any geometrical object should be considered as a set of points [4]. Also the course asserts that “the straight line can be defined by two points” and “the plane can be defined by three points which do not belong to one straight line”.



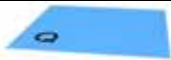
It will be more correct to begin a course of Descriptive Geometry with well-known geometry fundamentals from a course of the highest geometry, Gilbert's geometry for example [1].

3. INNOVATIVE CONTENT OF DISCIPLINE “DESCRIPTIVE GEOMETRY”

The teacher should teach that “point”, “straight line”, “plane” are the main geometrical objects, that dimension of a point equals three, that a

straight line – four, the plane – three in Euclid’s space . It is given in Table 1.

Table 1: Dimension of basic geometric objects in Euclid’s space E_3 .

Object	Illustration	Dimension
Point		3
Line		4
Plane		3

3.1 Basic notions

The straight line, the plane and space consist of points which, in turn are dot sets. However, the plane and space can be defined as sets of straight lines. Thus, the space can be defined as a set of the planes that is as plane space. These sets “are enclosed” into each other: the point – in a straight line; the point and the straight line – into a plane; the point, the straight line and the plane is objects of three-dimensional space E_3 . Such combination of sets can be presented in the form of a peculiar flag for which: the plane P is a panel which can be defined as a set of straight lines d ; the straight lines d is a flag-staff which in turn can be defined as a set of points A ; the point A is a star. It is shown in Figure 1. Visual similarity with the flag having a star, a flagstaff and a panel, gave such geometrical set the name “Flag” [3].

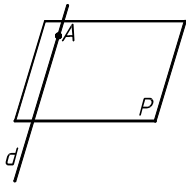




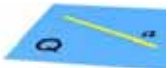



Figure 1: Graphic interpretation of the concept “Flag”.

The teacher has to tell about parameters defining geometrical objects, their dimension, the structural relations which will separate an ele-

ment from a set of elements. The main incidence conditions of geometrical objects are given in Table 2.

Table 2: Dimensions of incidence conditions of the main geometrical objects in Euclid’s space E_3 .

Incidence condition	Illustration	Dimension
<i>Point</i>		
The point belongs to a space		3
The point belongs to a straight line		2
The point belongs to a plane		1
<i>Straight line</i>		
The straight line belongs to a space		4
The straight line belongs to a plane		2
<i>Plane</i>		
The plane belongs to a space		3

We have defined the basic geometrical elements so now it is possible to consider the theme “Creation of Surfaces”.

4. INNOVATIVE TRAINING APPROACH TO THE THEME “CREATION OF SURFACES”

The world of surfaces is various and boundless. It lasts from the elementary plane with its mathematical severity, to the most difficult and freakish forms of curvilinear surfaces which do not give in to the exact mathematical description [4].

The surface is a continuous set of lines at which dimension is equal to unit. An all-encompassing uniform classification of sur-

faces does not exist to cover all types of surfaces. There are only private classifications which unite surfaces in groups only according to one characteristic. In Descriptive Geometry a creation of a surface is done graphically. Therefore, classification based on the line type creating a surface is more common.

In Descriptive Geometry students study a kinematic way of creation of surfaces. Surface is represented by only some positions of the line forming this surface on the drawing. The rule of movement of the forming line in space is shown in the form of the motionless line or set of lines. Lines which show the rule of movement of the forming line are called “Directing Lines”.

All classifications of surfaces are united by general dependence: if a dimension of a set of lines is equal to k in space, then a dimension of a set of forming lines of this surface is equal to a unit – a one-parametrical set of lines. We receive the rule of creation of surfaces.

To create a surface from a presented set of lines in space, it is necessary:

- 1) to impose such links on a set of lines for which the sum of dimensions is equal to $k-1$;
- 2) to provide a condition of compatibility of these links.

Thus, we have to know all possible links and their dimensions for the line which creates a surface.


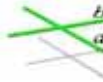




The new approach of teaching to the theme “Creation of surfaces” is based on the logical selection conditions that are imposed on the surface determinant. If the straight line is forming for a surface and its dimension in Euclid's space is equal to four, then the dimension's sum of conditions for creation of a line surface has to be equal to three. Indeed, the surface is a continuous set of lines in which dimension is equal to one. Therefore, to construct a surface we need to “tie” three of four parameters of imposed conditions on the line.

Conditions of an arrangement of a straight line forming a surface, relative to geometrical objects of Euclidean space, can be grouped ac-



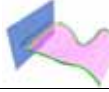


cording to the form of the objects participating in these conditions. We suggest to name the groups of conditions based on the first letter of the names of objects. For example, all conditions of the relations for a forming straight line with a point are called P1, P2, etc. from the word “Point”; the letter “L” (“Line”) is applied to straight; for curves – “C” (“Curve”); for the plane – “F” from “Flatness”; for a surface – “S” from “Surface” [2].

Each condition has a corresponding dimension in a three-dimensional space. For example, Table 3 conditions for a straight line that forms linear surfaces.

Table 3: Geometrical conditions imposed on the rectilinear forming line.

Name	Condition	Condition illustration	Dimension
P1	Passing through a point		2
L1	Straight line crossing		1
L3	In parallel to a straight line		2
L4	Perpendicularly to a straight line		1
C1	Curve crossing		1
F1	In parallel to the plane		1

It is possible to create various linear surfaces by carrying out sets of possible conditions. We have to remember that the sum of dimensions of the chosen conditions has to be equal to three. Examples of linear surfaces which can be created from conditions of Table 3 are presented in Table 4.

Table 4: Examples of linear surfaces.		
Set of conditions (dimension)	Surface determinant	Surface illustration
P1 • C1 (2+1)	- to pass through a point and to cross a curve	
L3 • C1 (2+1)	- parallel to a straight line and to cross a curve	
C1 • C1 • F1 (1+1+1)	- to cross two curves and parallel to the plane	
L1 • C1 • F1 (1+1+1)	- straight line crossing, crossing of a curve and parallel to the plane	
L1 • L4 • C1 (1+1+1)	- straight line crossing, perpendicular to a straight line and curve crossing - the screw line	

4.1 Electronic illustrator “Creation of surfaces”

It is very important for students to not only set surfaces theoretically, but also to see these surfaces, and the process of their creation. Therefore, we offer the electronic illustrator “Creation of surfaces” as the training manual. Students in a game form choose geometrical conditions from the offered list, which, when imposed on the rectilinear forming line create a linear surface.

5. CREATION OF CURVILINEAR SURFACES

Curvilinear surfaces can be studied analogously. For a curvilinear surface, the forming line is a curve. The curve which creates a surface can be constant or variable size. Quadric Curves (a


circle, parabolas, ellipses and hyperboles) are used as the curvilinear forming line.




The surface in three-dimensional space is an one-parametrical set of lines; therefore, while creating a surface it is important to know dimension of the line which creates this surface. So, the circle in space E_3 has dimension equals to six, a parabola – seven, and an ellipse and a hyperbole – eight.

If we follow the logic while creating linear surfaces, it is necessary to impose conditions on the curvilinear forming line for which the sum of dimensions would be one unit less, than dimension of the forming line. That is for a circle – five, for a parabola – six, for an ellipse and a hyperbole – seven.

However, at formation of a curvilinear surface it is necessary to consider that curves which are selected as forming (circles, parabolas, ellipses and hyperboles) are flat curves for which all its points belong to a plane. We know that dimension of a plane in E_3 space is equal to two. Then for conditions of the curvilinear forming line the sum of dimensions has to be equal: for circles – three; for a parabola – four; for an ellipse and a hyperbole – five.


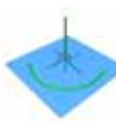
For convenience we divided conditions for curvilinear forming lines into two groups and gave names: the letter “C” (from “Curve”) for the forming line and the letter “F” (from “Flatness” - “plane”) for the plane in which the forming line is located. The second letter in designation will designate the objects entering the relations with the forming line, to similarly accepted designations. Some geometrical conditions for the curvilinear forming line are given in Table 5.

Table 5: Conditions for the curvilinear forming line.			
Name	Condition	Condition illustration	Dimension
CL1	Crossing with a straight line		1

CL2	Crossing with a straight line in a certain point		2
CC1	Crossing with a curve		1
CS1	Touch with a surface		1

Some conditions for the plane of the curvilinear forming line are given in Table 6.

Table 6: Conditions for the plane of the curvilinear forming line.




Name	Condition	Condition illustration	Dimension
FL1	The plane of the curvilinear forming line passes through a straight line		2
FL2	The plane of the curvilinear forming line is perpendicular a straight line		2

We can make combinations of these conditions to construct nonlinear surfaces. Thus some conditions can be multiple.

We should not forget dependence of conditions for formation of curvilinear surfaces: $(\Sigma C)[(\Sigma F)]$. Where "C" – conditions for the forming line, "F" – for the plane of the forming line. The sum of dimensions of conditions (ΣC) и (ΣF) for each corresponding view of forming lines will have the certain value: $(\Sigma C) = 3$; $(\Sigma F) = 2$ – for circles; $(\Sigma C) = 4$;

$(\Sigma F) = 2$ – for parabolas; $(\Sigma C) = 5$; $(\Sigma F) = 2$ – for ellipses and hyperboles. Some examples of curvilinear surfaces are given in Table 7.

Table 7: Examples of curvilinear surfaces.

Set of conditions (dimension)	Surface determinant	Surface illustration
<i>The forming line - a circle</i>		
CL1 • CC1 • FL2 $(\Sigma C) = 1 + 1 + 1$ dimension is a change of radius of the forming line; $(\Sigma F) = 2$	Crosses a straight line, crosses a curve and the plane of the forming line is located perpendicular to this straight line.	
<i>The forming line - a parabola</i>		
CL2 • CC1 ² • FL1 $(\Sigma C) = 2 + (1 + 1);$ $(\Sigma F) = 2$	Crosses a straight line at a certain point, crosses a parabola and the plane of the forming line passes through this straight line.	
<i>The forming line - an ellipse</i>		
CL2 ² • CS1 • FL1 $(\Sigma C) = (2 + 2) + 1;$ $(\Sigma F) = 2$	Crosses a straight line at two certain points, the touch with a surface and the plane of the forming line passes through this straight line.	
<i>The forming line - a hyperbole</i>		
CL1 ⁴ • FL2	Crosses four parallel straight	

$(\Sigma C) = 1 + 1 +$ lines and the
 $+ 1 + 1$ planes of the
 $+ 1$ dimension forming line is
 is a parallel- located perpen-
 ism of the dicularly to this
 straight lines; straight line.
 $(\Sigma F) = 2$



6. CONCLUSIONS

In this paper we offered a logical-constructive training approach to the theme "Creation of Surfaces" which allows students develop scientific research skills and to provoke interest in construction and research of difficult form surfaces. This approach can be used in scientific research work with the students and as they further their training in postgraduate studies.

REFERENCES-ССЫЛКИ

- [1] D. Hilbert. Foundations of geometry: [translation from German]. *Seyatel publishing house, Leningrad, Russia*, 1923.
- [2] N.V. Kaygorodtseva. Innovative methodology of descriptive geometry: monograph. *Omsk state technical university, Russia*, 2013.
- [3] B.A. Rosenfeld. Non-Euclidean spaces. *Nauka publishing house, Moscow, Russia*, 1969.
- [4] S.A. Frolov. Descriptive geometry. *Mashinostroenie publishing house, Moscow, Russia*, 1983.

ABOUT THE AUTHORS

- 1. Natalya V. Kaygorodtseva, PhD of pedagogical sciences, associate professor of the department "Engineering geometry and CAD", Omsk state technical university
- 2. Tatyana N. Kaygorodtseva, student of the radio-technical faculty, Omsk state technical university

MAP PROJECTION OF THE WORLD MAP BY LEONARDO DA VINCI

Miljenko LAPAINE

University of Zagreb, Croatia

ABSTRACT: Leonardo da Vinci produced a world map about 500 years ago. The map consists of eight parts of equal size, each of which is a projection of one eighth of a globe into a plane. Lines separating those parts are the equator and two arcs of meridian circles. Those parts are equilateral triangles, also known as Reuleaux's triangles after German engineer Franz Reuleaux who lived in the 19th century. Each side of those triangles is an arc of a circle with opposite vertex as the centre of the circle. Diameters of those triangles equal five inches, making the scale of the world map very small. Meridians and parallels are not shown on the map, which is not unusual because it is from a period in which maps were produced to show discoveries rather than help sailors on their voyages. The da Vinci map projection is unique, without predecessors in Antique or Middle Ages. It is unknown which map projection was applied by da Vinci. However, the paper demonstrates the following:

1. The angle in the da Vinci-Reuleaux triangle equals 120° .
2. Da Vinci's projection is not an oblique aspect of azimuthal projection (neither conformal, nor equal-area, nor equidistant).
3. It is possible to derive equations of a pseudoconic map projection with the following properties: parallels of latitude are arcs of concentric circles with the centre in the pole; meridians of longitude are arcs of circles with the centre on the coordinate axis. Parallels in the projection plane divide the central meridian into equal parts.

Keywords: Leonardo da Vinci, world map, Reuleaux triangle, map projection.

1. INTRODUCTION

The map consists of two sheets. Recto: Map of the northern hemisphere in four segments. Verso: Map of the southern hemisphere in four segments (Figure 1). The map is preserved in the Royal Collection of Her Majesty Queen Elizabeth II and bears the signature RCIN 991393. According to official data [17], the map has been preserved in the Royal Collection since 1690, likely thanks to Charles II, King of Great Britain. The map was inherited from Charles II by Francesco Melzi, from whose heirs it has been purchased by Pompeo Leoni, c. 1582–90, and Thomas Howard, 2nd Earl of Arundel, by 1630. The size of each map sheet is 27.7×28.8 cm.

Major [18] was the first to publish da Vinci's map and describe it in detail. He described all places on da Vinci's map in great detail, but he did not analyse the map's projection. He only noted that graticule was not shown on the map. According to Major [18], da Vinci's world map has the following exceptional properties: 1. At the time, it was the earliest known world map with the name America on it; 2. It was the earliest map showing that the Western coast of America is not connected to Asia; 3. At the time, it was the only known map indicating the earlier notion of a great Southern continent, prior to the discovery of the Straits of Magellan, after which that assumption was considered to be true and drawn on maps as an indefinite

continuation of the then discovered Tierra del Fuego. See also [12, 23, 24].

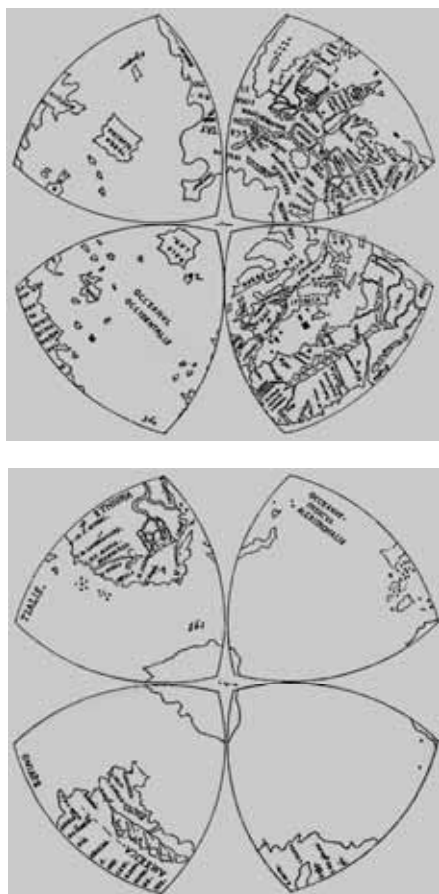


Figure 1: Recto: Map of the northern hemisphere in four segments. Verso: Map of the southern hemisphere in four segments. Royal Collection of Her Majesty Queen Elizabeth II, Windsor, RCIN 991393

Nowadays, it is known that the first map with the name America on it is somewhat older than da Vinci's world map. It is *Universalis Cosmographia* by Martin Waldseemüller, which was a printed wall map first published in

April 1507. Nevertheless, this fact does not diminish the value of da Vinci's map, which still bears some unresolved issues.

Also worth mentioning is that Major [18] knew of Martin Waldseemüller and his work, but he was obviously unaware of the 1507 map, considering he gave priority to da Vinci.

1.1 Authorship and map's dating

Authors who have written about da Vinci's map do not agree on the authorship of the drawing in the da Vinci legacy. Some consider it a copy of the original by da Vinci [19, 20]. Popham and Wilde [21] doubt the attribution to Leonardo; Clark and Pedretti [5] suggest the possible attribution to Francesco Melzi.

Francesco Melzi (fl. 1491 – 1568/1570) was the son of a Milanese noble family. Melzi joined the household of Leonardo da Vinci in 1506. He accompanied Leonardo on trips to Rome in 1513 and to France in 1517. As a painter, Melzi worked closely with and for Leonardo. Some works which were attributed to Leonardo during the nineteenth century are nowadays ascribed to Melzi. Upon Leonardo's death, Melzi inherited Leonardo's artistic and scientific works, manuscripts, and collections and would faithfully administer the estate. Melzi wrote to Leonardo's brothers to notify them of his death, and in this letter he described Leonardo's love for his pupils as "*sviscerato e ardentissimo amore*", a selfless and incandescent love. Returning to Italy, Melzi married and fathered a son, Orazio. When Orazio died on his estate in Vaprio d'Adda, his heirs sold the collection of Leonardo's works [10].

Major [18] estimated the map was produced in 1513, while others consider it was probably produced in 1514 or 1515. According to recent research, it was produced around 1550 [17].

2. MAP DESCRIPTION FROM THE MAP PROJECTION POINT OF VIEW

The map consists of two sheets, each representing 4 octants, i.e. each octant is a projection of one eighth of a globe into a plane. Lines di-

viding these parts are the equator and two arcs of meridian circles. Of course, these parts are equilateral triangles known as Reuleaux triangles, after German engineer Franz Reuleaux, who lived in the 19th century. Each side of those triangles is an arc of a circle with the centre in the opposite vertex of the triangle. Diameters of the triangles equal exactly five inches, making the map scale very small. The map does not show meridians and parallels, which is not unusual since maps produced at the time were meant to represent discoveries and not help sailors on their journeys.

Da Vinci's world map was first published and described in detail by Major [18]. He described all the places represented in the map, but did not analyse its projection.

2.1 Da Vinci-Reuleaux Triangles

Da Vinci, a versatile scientist and artist, was also interested in regular polygons. He was especially interested in constructing regular polygons using only a compass. Thirteen of his mirror-written manuscript volumes have been preserved. These volumes were published at the end of the 19th century and contain several constructions of regular pentagons, octagons, nonagons and octadecagons. This was analysed by Cantor [4], who concluded that only one of the constructions is correct, another is completely unusable, while others contain only minor errors and could be used in practice.

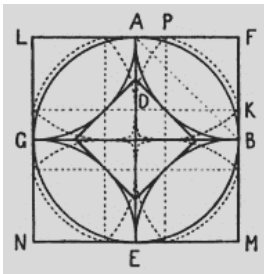


Figure 2. Uhden's attempt of explaining the construction of da Vinci's triangles [25]

Fiorini [9] was the first to research the mathematical foundation of da Vinci's world map and described a possible way of constructing da Vinci's curvilinear triangles, which represent one spherical octant in the projections. Fiorini also determined the relation between the size of the spherical octant on a sphere and in the projection, as well as the map's potential projection. Fiorini explains the construction of da Vinci's triangles in the following way:

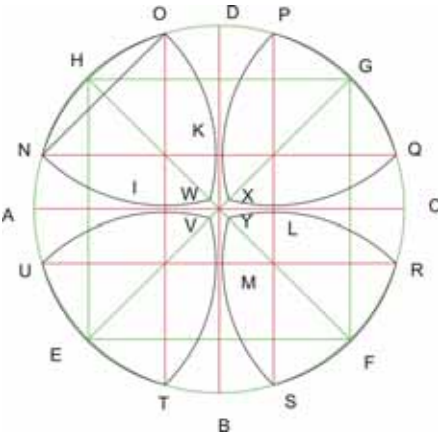


Figure 3. Reconstruction of an unsuccessful attempt at explaining the construction of da Vinci's triangles according to Keuning [15]

If we wish to draw four sectors of a hemisphere using circle arcs with radius R so that neighbouring sectors touch, we have to draw two perpendicular straight lines which will represent their central meridians, then determine points on them which are $R\sqrt{2} \square R$ away from their intersection, followed by drawing a line from each of those points, left and right under the angle of 45° with a length of $R/\sqrt{2}$. Endpoints of these lines are endpoints of four equatorial arcs which we can use to construct the four sectors.

Uhden [25] describes one possible construction of da Vinci's triangles according to Fiorini [9] and also proposes one solution of con-

structing da Vinci's triangles (Figure 2) based on a sketch by da Vinci from *Codice Atlantico* (fol. 156r) [6].

Keuning [16] incorrectly describes the construction of da Vinci's triangles (Figure 3).

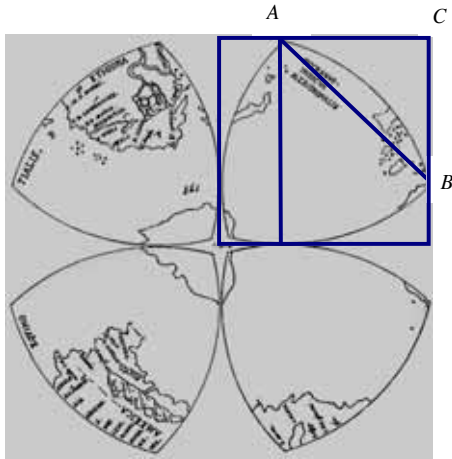


Figure 4. Length of a side of the square is equal to the length of a side of the equilateral triangle inscribed in da Vinci's curvilinear triangle

It is somewhat unusual the mentioned authors were unable find out the very simple construction (Fig. 4). Namely, triangle ABC is isosceles and rectangular with the base AB equal to the length of a side of the square, with $AC = BC$ and equal to half the length of the square diagonal.

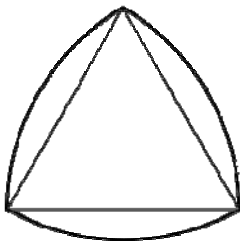


Figure 5. Reuleaux's triangle, which was used by da Vinci to construct his world map [22]

Reuleaux's triangle is a curve with constant width based on an equilateral triangle. The distance between each point on a side from the opposite vertex is the same (Fig. 5). Reuleaux's triangle is the simplest and most famous among Reuleaux's polygons. The triangle was named after German engineer Franz Reuleaux, who lived in the 19th century. Considering Reuleaux lived four centuries after da Vinci, I propose the triangle be called the da Vinci-Reuleaux triangle.

3. PROJECTION OF DA VINCI'S WORLD MAP

No written record exists about the projection used for da Vinci's world map. The lack of a graticule on the map poses an additional problem. Therefore, it is not possible to determine with certainty how the map was drawn, i.e. what equations could describe the projection the map was produced in.

Fiorini [9] was the first to research the mathematical foundation of da Vinci's world map and described one way of constructing da Vinci's curvilinear triangles (see chapter 2.1). In addition, Fiorini described one projection, i.e. a method of representing the graticule in da Vinci's triangles according to O. Finé, who published his method in 1549, 30 years after da Vinci died.

Addressing the first Le Testu projection, Anthiaume [1,2] claims it is the da Vinci projection. However, he did not prove his claim. Furthermore, Anthiaume states the same projection is featured in O. Finé's *De Mundi Sphaera sive Cosmographia libri V*, which is incorrect. Meridians are arcs of circles in Finé's projection, which is not the case for Le Testu projection [1].

According to Uhden [25], the da Vinci projection is considerably different from all existing projections. It is a fact that da Vinci was familiar with the issue of mapping a sphere into a plane [13]. This can be concluded from one of his sheets from *Codice Atlantico* (fol. 191r), which includes drawings of orthographic nor-

mal and transversal projection, oval projection by P. Apian and several drawings of a ball and its parts.

Keuning [16] considers the da Vinci projection unique, without predecessors in Antiquity or Middle Ages.

3.1 Oronce Finé and his construction of octant projections

Oronce Finé (in Latin, Orontius Finaeus) (1494–1555) was a French mathematician and cartographer. In 1531, he was appointed to the chair of mathematics at the Collège Royal (the present Collège de France), founded by Francis I of France, where he taught until his death. Although primarily a populariser, Finé was one of the most prolific authors of mathematical books of his age. He worked in a wide range of mathematical fields, including practical geometry, arithmetic, optics, gnomonics, astronomy, and instrumentalism. In *De rebus mathematicis* (1556), he provided the value of π in the form of $3 \frac{11}{78} = 3.1410$.

In 1542, Finé published *De mundi sphaera* (On the Heavenly Spheres), a popular astronomy textbook whose woodcut illustrations were much appreciated. His writing on astronomy included guides to the use of astronomical equipment and methods (e.g. the ancient practice of determining longitude by coordinated observation of lunar eclipses from two fixed points with enough distance between them to make the phenomena appear at different times of night.) He also described more recent innovations, such as an instrument he called a *méthéoroscope* (a modified astrolabe).

He complemented his explanatory work with direct contributions. His woodcut map of France (1525) is one of the first of its kind. He constructed an ivory sundial in 1524 which still exists.

Finé's heart-shaped (cordiform) map projection may be his most famous illustration, and it was frequently employed by other notable cartographers, including Peter Apian and Gerardus Mercator.

Finé attempted to reconcile discoveries in the New World with old medieval legends and information (derived from Ptolemy) regarding the Orient. Thus, on one of his two world maps, *Nova Universi Orbis Descriptio* (1531), the legend marked *Asia* covers both North America and Asia, represented as one landmass. He used the toponym "America" for South America, and thus Marco Polo's *Mangi*, *Tangut*, and *Catay* appear on the shores of the present-day Gulf of Mexico. On the same map, Finé drew *Terra Australis* to the south, including the legend "recently discovered but not yet completely explored," by which he meant the discovery of Tierra del Fuego by Ferdinand Magellan.

Finé is the author of *Le sphere du monde, proprement dite Cosmographie*, ... [8]. It consists of five books in French and there is a manuscript from 1549. Fiorini states 1551, while it was printed in Latin in 1551 as *De Mundi Sphaera, sive Cosmographia*, lib. V, ... [7]. The seventh chapter in the fifth book is titled in French as *Comment la huitième et quarte partie, et la moitié du globe terrestre peuvent être commodément réduites en plate forme*, in Latin as *Qua ratione octava, seu quarta pars, atque dimidia terrestres orbis contestura in plano commodissime delineare possit* and in English as *The most suitable way of drawing an eighth, a fourth or a half of the Earth's sphere in a plane*.

Of mention is that Anthiaume [2] stated that O. Finé published *De Mundi Sphaera sive Cosmographia libri V* in Paris in 1542 (*Parisiis, ex officina Simonis Colinoei*, in-folio) and that Finé translated it to French and published the translation in Paris in 1522 (*La Sphere du Monde, proprement dite Cosmographie*, 64 fol. in-4°). In addition, the Houghton Library at Harvard preserves Finé's manuscript in French from about 1549 missing a part of the seventh chapter from the fifth book, and which exists in the Latin edition from 1542. Nordenskiöld [19] indicates *Sphaera Mundi*, Lutetiae Parisiorum by Orontius Finaeus from 1551. It is clear the work by Finé was published in several editions

in French and Latin in the middle of the 16th century.

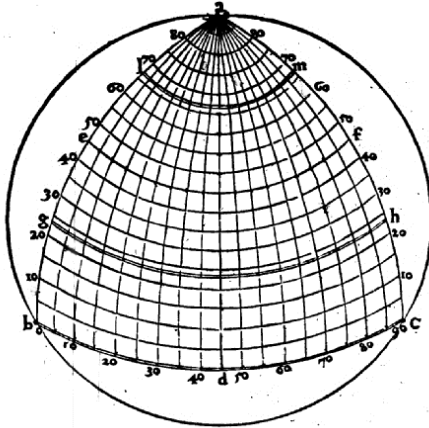


Figure 6. Drawing from Finé's book [7]

This is how Finé describes map construction in the projections: "If one wishes to draw one eighth of the Earth's sphere in a plane encompassed with circles from the equator to one of the poles, it is first necessary to draw a circle with an arbitrary size and divide its circumference into three equal parts. By using setting one part of a compass on any of the three parts and the other part on one of the other two, one has to form a triangle with curved and equal sides. One of its sides is going to represent one fourth of an equator, and the other two sides are going to be images of marginal meridians. Then, a fourth of the equator image is divided into two parts and a straight line is drawn from the pole image to the division point so that is an image of the central meridian. After that, the straight line has to be divided into 90 equal parts and draw fourths of parallels through individual points, together with the northern and southern tropic and Arctic or Antarctic polar circles. In addition, the equator and other central parallels have to be divided into 90 equal parts and draw parts of other meridians from the pole through those points. After marking numbers of degrees of latitude and longitude, it is necessary to draw

individual places of the selected eighth" [7,8] (Fig. 6).

After this, Finé continues by describing and illustrating projections of one fourth and one half of the Earth's sphere, northern or southern, as an expansion of the procedure for one eighth of the Earth's sphere. This part does not exist in the French manuscript [8], but it is feature in the Latin version [7]. Finally, Finé provides a description and illustration for a projection of a hemisphere representing both poles.

3.2 Angle in the da Vinci-Reuleaux triangle

In Figure 5, an angle in the da Vinci-Reuleaux triangle clearly equals 120°. Namely, the angle consists of an angle in the centre of a regular triangle (60°) and two side angles, each of which equals 90°-60°=30° because a side of a regular triangle is at the same time a radius of a side of the da Vinci-Reuleaux triangle. The same result can also be reached in other simple ways.

3.3 Da Vinci projection is neither one of known oblique azimuthal projections

Based on the map projection theory and experience, one could assume that one octant in the da Vinci projection could be in one of the azimuthal projections. Using the result from the previous chapter, it is going to be proved the da Vinci projection is neither conformal, nor equal-area, nor equidistant azimuthal.

Equations of azimuthal projections are usually written as follows in Equation (1):

$$x = \varphi(z) \cos \varphi, \quad y = \varphi(z) \sin \varphi \quad (1)$$

where the relation between spherical coordinates φ and z and geographic coordinates φ and λ is described by following equations (2):

$$\begin{aligned} \sin z \cos \varphi &= \sin \varphi_0 \cos \varphi_0 \sin \varphi_0 \cos \varphi_0 \cos(\varphi - \varphi_0) \\ \sin z \sin \varphi &= \sin(\varphi - \varphi_0) \cos \varphi_0 \\ \cos z &= \sin \varphi_0 \sin \varphi + \cos \varphi_0 \cos \varphi \cos(\varphi - \varphi_0), \end{aligned} \quad (2)$$

where φ_0 and λ_0 are given coordinates of the pole of spherical coordinates in the system of geographic coordinates. Without reducing generality, we can assume that $\lambda_0 = 0$. Let us assume the projection is an oblique azimuthal projection of the first octant and that $\varphi_0 = \varphi/4$. Therefore, the coordinates of the central point of the mapped area are $\varphi_0 = \varphi/4, \lambda_0 = 0$.

Let β be the angle which the tangent on meridian $\varphi = \varphi/4$ closes in the North Pole $\varphi = \varphi/2$ with the y axis. Using formulae (2), we get that coordinates of the North Pole are $\varphi = 0$ and $z = \varphi/4$. In general,

$$\begin{aligned} \tan \varphi &= \frac{dx}{dy} = \frac{\frac{d\varphi}{dz} \frac{dz}{d\varphi}}{\frac{d\lambda}{dz} \frac{dz}{d\lambda}} = \frac{\frac{d\varphi}{dz} \cos \varphi \sin \varphi}{\frac{d\lambda}{dz} \sin \varphi + \varphi \cos \varphi} = \\ &= \frac{\frac{d\varphi}{dz} \frac{dz}{d\varphi} \cos \varphi \sin \varphi}{\frac{d\varphi}{dz} \frac{dz}{d\varphi} \sin \varphi + \varphi \cos \varphi} \end{aligned} \quad (3)$$

The derivation of the third equation in (2) by φ gives

$$\varphi \sin z \frac{dz}{d\varphi} = \sin \varphi_0 \cos \varphi \cos \varphi_0 \sin \varphi \cos \varphi \quad (4)$$

from where $\frac{dz}{d\varphi}$ is obtained easily. By deriving the second equation from (2) by φ ,

$$\cos z \frac{dz}{d\varphi} \sin \varphi + \sin z \cos \varphi \frac{d\varphi}{dz} = \varphi \sin \varphi \sin \varphi \quad (5)$$

from where $\frac{d\varphi}{dz}$ is obtained easily. In the North Pole, we have $\frac{dz}{d\varphi} = \sqrt{2}/2$ and $\varphi/\varphi = 1$. By substituting these values into (3), we get

$$\tan \varphi = \frac{\sqrt{2}}{2} \frac{1}{\varphi} \frac{d\varphi}{dz} \quad (6)$$

For conformal projections,

$$\varphi = 2R \tan \frac{z}{2} \quad (7)$$

so $\frac{d\varphi}{dz} = \frac{R}{\cos^2 \frac{z}{2}}$ and $\tan \varphi = 1$ in the ob-

served point. Considering that $\tan 60^\circ = \sqrt{3}$, we conclude the projection is not the conformal azimuthal projection.

For equal-area azimuthal projections,

$$\varphi = 2R \sin \frac{z}{2} \quad (8)$$

so $\frac{d\varphi}{dz} = R \cos \frac{z}{2}$ and $\tan \varphi = \frac{\sqrt{2}}{4} \cot \frac{\varphi}{8} \approx 0.8$ in the observed point. We conclude the projection is not the equal-area azimuthal projection.

For azimuthal projections which are equidistant along meridians,

$$\varphi = Rz \quad (9)$$

so $\frac{d\varphi}{dz} = R$ and $\tan \varphi = \frac{2\sqrt{2}}{\varphi} \approx 0.9$ in the observed point. We conclude the projection is not the equidistant azimuthal projection.

The natural question to rise is whether there is an oblique azimuthal projection for which (6) would be true and $\varphi = 60^\circ$? In other words, let us look for the solution of the differential equation

$$\frac{\sqrt{2}}{2} \frac{1}{\varphi} \frac{d\varphi}{dz} = \sqrt{3} \quad (10)$$

By separating variables, we easily find the solution

$$\varphi = Ke^{\sqrt{6}z}, \quad (11)$$

where K is the integration constant. Considering that for $z = 0$, it should be $\varphi = 0$, we conclude da Vinci's map of one octant was not produced in any oblique azimuthal projection.

3.4 Possible projection of da Vinci's map

Let us set a rectangular coordinate system x, y so that y axis is horizontal and in that system define an equilateral triangle with vertices in $(0, \varphi a/2)$, $(0, a/2)$ and $(a\sqrt{3}/2, 0)$. Let parallels be arcs of circles with centre in the image of North Pole, i.e. the point $(a\sqrt{3}/2, 0)$ and radius

$$\varphi(\lambda) = a \left(1 - \frac{2\lambda}{\pi} \right). \quad (12)$$

Equation of parallels is going to be

$$\left(x - \frac{a}{2}\sqrt{3} \right)^2 + y^2 = \varphi^2(\lambda). \quad (13)$$

Let meridians also be arcs of circles with centre on y axis in the point with coordinates $\left(0, \sqrt{\varphi^2(\lambda) - 3a^2/4} \right)$, where

$$\varphi(\lambda) = a \left(\frac{3\lambda}{16\pi} + \frac{|\lambda|}{\pi} \right) \quad (14)$$

is the radius of the meridian. Therefore, the equation of the meridians is

$$x^2 + \left(y - \sqrt{\varphi^2(\lambda) - 3a^2/4} \right)^2 = \varphi^2(\lambda). \quad (15)$$

It is easily seen that parallels in the projection divide the central meridian in the projection into equal parts. Considering that parallels in the projection are arcs of concentric circles, and meridian arcs of circles symmetrical to the central meridian which is mapped as a section of a straight line, the projection is a pseudoconic projection. General equations of those projections are

$$x = q \varphi(\lambda) \cos \varphi(\lambda), \quad y = \varphi(\lambda) \sin \varphi(\lambda). \quad (16)$$

In our case, it can be shown that

$$q = a\sqrt{3}/2, \quad (17)$$

and $\varphi = \varphi(\lambda)$ should be the solution of the projective equation

$$q \cos \varphi \sqrt{\varphi^2(\lambda) - q^2} \sin \varphi = \frac{\varphi(\lambda)}{2}. \quad (18)$$

Therefore, in order to calculate rectangular coordinates x, y in the projection plane for given values of geographic coordinates λ and φ , it is first necessary to solve the equation (18), taking into consideration (12) and (14), and then use (16).

4. AFTER DA VINCI'S WORLD MAP

Fiorini [9] mentioned navigational atlas by Francesco Gisolfo (Sign. 3140) from the end of the 16th or beginning of the 17th century, which

is preserved in the Biblioteca Riccardiana u Florence. The world map in the atlas consists of eight da Vinci-Reuleaux triangles. However, those triangles are distributed in two rows as segments of globe, in contrast to da Vinci. A similar representation can be found in two other navigational atlases by unknown authors, probably from the end of the 16th century. One is preserved in the National Library in Naples (Sign. VIII.D.6) and the other in the University Library in Genova (Sign. G.V.32).

An example of such a map is the world map by Daniel Angelocrator and engraved by Nicolaas Geelkercken in 1616 [15].

Obviously, the da Vinci projection is not widely applied in practice. In 1914, the Cahill projection [3,11] appeared, which divided the globe into octants, mapped them into a plane, resulting in a butterfly shaped map. The Cahill projection continues to inspire cartographic researchers [14].

5. CONCLUSION

The world map drawn by Leonardo da Vinci or Francesco Melzi consists of eight equal parts, each of which is a projection of one eighth of a globe into a plane. Those parts are equilateral triangles, also known as Reuleaux's triangles, after German engineer Franz Reuleaux. Considering Reuleaux lived four centuries after da Vinci, I propose the triangle be called the da Vinci-Reuleaux triangle.

The paper shows how some authors struggled with explaining construction of da Vinci-Reuleaux triangles on da Vinci's world map, even though the construction can be explained easily.

Considering that da Vinci did not leave a written record of the projection in which he produced his world map, the final part of the paper addresses that issue. It is shown the map was not produced in an azimuthal projection, and a possible pseudoconic projection is described in which parallels are arcs of concentric circles, while meridians are arcs of circles with centres on a straight line.

ACKNOWLEDGMENTS

I would like to thank Mira Miletic Drder from the Map and Atlas Collection of the National and University Library in Zagreb, Renata Šolar from the Map and Graphic Collection of the National and University Library in Ljubljana and Patrizia Licini for their help in acquiring literature. I would also like to thank Vida Zadelj Martić from the Faculty of Geodesy of the University of Zagreb for translating Fiorini's article from Italian and Olga Perić for translating a chapter of Finaeus's book from Latin. In addition, I would like to thank Rosie Razzall, Curator of Prints and Drawings, Royal Collection Trust, Windsor Castle for her help related to the issue of da Vinci's map authorship.

REFERENCES

- [1] Anthiaume, A. M. F. (1911): Un pilote et cartographe havrais au XVIIe siècle: Guillaume Le Testu, *Bulletin de Géographie Historique et Descriptive*, Paris, Nos 1-2.
- [2] Anthiaume, l'abbé A. M. F. (1916, 1919): Cartes marines, Constructions navales, Voyages de découverte chez les Normands 1500-1650, vol. I, II, Paris.
- [3] Cahill, B. J. S. (1937): Die Schemetterlingskarte, *Peterm. Mitt.* 129 i dalje.
- [4] Cantor, M. (1890): Über einige Konstruktionen von Lionardno da Vinci, Mitteilungen der Mathematischen Gesellschaft in Hamburg, Band 2 – zugleich Festschrift herausgegeben von der Mathematischen Gesellschaft in Hamburg anlässlich 200jährigen Jubelfestes, 2. Teil, Wissenschaftliche Abhandlungen, 8–15.
- [5] Clark, K., Pedretti, C. (1935): The Drawings of Leonardo da Vinci in the Collection of Her Majesty The Queen at Windsor Castle, nos. 01393 and 01393 bis. Phaidon Press Ltd, 2nd edition 1968.
- [6] Da Vinci, L. (1891–1904): Il Codice Atlantico nella Biblioteca Ambrosiana di Milano, riprodotto e pubblicato dalla Reale Accademia dei Lincei etc., Milano.
- [7] Finaeus, O. (1555): De mundi sphaera, sive cosmographia, libri V, M. Vascosanum, Paris. Source : Bibliothèque nationale de France, <http://catalogue.bnf.fr/ark:/12148/cb37229968c>
- [8] Finé, O. (1549): Le sphere de monde : proprement dicte Cosmographie : manuscript, MS Typ 57. Houghton Library, Harvard University, Cambridge, Mass., <http://nrs.harvard.edu/urn-3:FHCL.HOUGH:4435151>
- [9] Fiorini, M. (1894): Il mappamondo di Leonardo da Vinci ed altre consimile mappe: *Rivista Geografica Italiana*, Roma, vol. 1, 213–223. [Reprinted in *Acta Cartographica*, 1969, vol. 4, 188–198.]
- [10] Francesco Melzi, http://en.wikipedia.org/wiki/Francesco_Melzi
- [11] Hammer, E. (1914): Die neue Weltkartenprojektion von Cahill, *Peterm. Mitt.* I, 323.
- [12] Henning, R. (1948): The Representation on Maps of the Magalhaes Straits before their Discovery, *Imago Mundi*, vol. V, 1948, 33–37.
- [13] Kemp, M. (2006): Leonardo da Vinci, The Marvellous Works of Nature and Man, Oxford, University Press.
- [14] Kerkovits, K., Györfy, J., Gede, M. (2013): Renewing Cahill's equal-area butterfly projection, Proceedings of the 26th International Cartographic Conference, Dresden. http://icaci.org/files/documents/ICC_proceedings/ICC2013/
- [15] Keuning, J. (1954): Nicolaas Geelkerken, *Imago Mundi*, XI, 174–177.
- [16] Keuning, J. (1955): The history of geographical map projections until 1600, *Imago Mundi*, XII, 1-24.
- [17] Leonardo da Vinci's map in the Royal Collection, <http://www.royalcollection.org.uk/collection/991393/recto-map-of-the-northern-hemisphere-in-four-segments-verso-map-of-the-southern>

- [18] Major, R. H. (1865): Memoir on a mappemonde by Leonardo da Vinci, *Archaeologia*, vol. XL, London.
- [19] Nordenskiöld, A. E. (1889): Facsimile-Atlas to the early history of cartography.
- [20] Oberhummer, E. (1909): Leonardno da Vinci and the art of the Rennaissance in its relation to geography, *The Geographical Journal*, vol. XXXII ili 33, 540–569.
- [21] Popham, A. E., Wilde, J. (1949): The Italian Drawings of the XV and XVI Centuries in the Collection of Her Majesty The Queen at Windsor Castle, no. 1193. Reprinted by Johnson Reprint Corporation, 1984, with Appendix by R. Wood.
- [22] Reuleaux triangle, http://en.wikipedia.org/wiki/Reuleaux_triangle
- [23] Snyder, J. P. (1993): *Flattening the Earth, Two Thousand Years of Map Projections*, The University of Chicago Press.
- [24] True, D. O. (1954): Some early maps relating Florida, *Imago Mundi*, XI, 73–84.
- [25] Uhden, R. (1938): Die Echtheit der Weltkarte des Leonardo da Vinci, *Comptes rendus d. Congr. intern. d. Géogr.*, Amsterdam, Tome II, 48–55.

ABOUT THE AUTHOR

Miljenko Lapaine studied Mathematics and graduated from the Faculty of Science, University of Zagreb, in the field of Theoretical Mathematics. He completed the postgraduate studies in Geodesy, in the field of Cartography at the Faculty of Geodesy in Zagreb by defending his Master's thesis *A Modern Approach to Map Projections*. He obtained his PhD from the same Faculty with a dissertation entitled *Mapping in the Theory of Map Projections*. He has been a full professor since 2003. He has published more than 800 papers, several textbooks and monographs. Prof. Lapaine is the chairman of the ICA Commission on Map Projections, a founder and president of the Croatian Cartographic Society and the executive editor of the *Cartography and Geoinformation* journal.

ON METHOD OF PRODUCING A DEFORMED SMALL FIGURE BASED ON AN ACTUAL PERSON'S BODY

Satoshi CHO, Hisashi SATO and Hideki YOSHII

Kanagawa Institute of Technology, Japan

ABSTRACT: The development of 3D printer and 3D scanner has a long history. Many of these products have been expensive. Recently, some companies have cheap 3D printer lineup. Producing a 3D object from 3D shape data is getting easier. Microsoft launched Kinect. Kinect is a motion sensing input devices for the Xbox and PCs. Kinect guesses the positions of 48 skeletal points on a human body. Kinect can behave like a 3D scanner. Kinect is not expensive. Using Kinect and the 3D printer leads to the development of a device that measures the 3D shape of a human and produces the 3D model.

In this paper, we propose a method of obtaining the 3D shape data of a person in the scene of depth information obtained with the use of Kinect, and generating a deformed 3D shape of the person based on the polygon data. Based on the skeleton information of the person obtained from Kinect, the method estimate body parts that contain the obtained polygonal information, and designs a grid for free form deformation based on this estimation. In this paper, we describe a method of making a deformed figure of the person, by outputting such deformed data with the use of a 3D printer.

Keywords: Geometry Processing, Deformed Figure, KINECT, 3D Printer.

1. INTRODUCTION

Many people including adults and children like anime and manga in Japan. Against this background, many studies are being conducted to develop systems to pull oneself into the world of manga and anime. For example, Touchi et al. developed a system that displays real-time *mappu* (unique symbols used for manga) for face images displayed in video images called OTAKU Vision[1]. Recently, Nara et al. developed a system called Manga Generator that converts photographed images into manga-like images to generate one-page manga[2], and displays this system for a long time in expositions such as "Manga Kingdom Tottori 'Manga Haku-Otsu'" and Japan Expo 2014 in Paris.

One of the things that fans of manga and anime like is so-called deformed figures. Deformed figures are figures that have a large head and a small body as shown in Figure 1. Deformed figures of characters of anime, manga, and games are on sale and have a large

market.

In this paper, we introduce a preliminary experiment needed for making a deformed figure of a real person by obtaining the three dimensional shape of a human body with the use of Kinect and deforming it.



Figure 1: Example of a deformed figure
(Made in 2012, Made by Satoshi CHO)

2. SHAPE AND DIMENSION OF ACTUAL DEFORMED FIGURES

A deformed figure comprises large head, small torso, short arms and short legs. In general, a deformed figure strikes a future pose such as stretching out its arm or raising its arm. The height of deformed figures is generally 80 to 100mm. Chart 1 shows the measured lengths of the head, torso, arm and leg parts of some deformed figures. There are no large differences among the lengths of the head, torso, and leg parts. However, there are wide variations in terms of the lengths of the left arm parts. This is supposedly because the left arms need to have characteristic shapes in order for the original characters to strike some kind of pose.

Table 1: Dimension of Deformed Figures.
(unit : mm)

head	torso	right arm	left arm	right leg	left leg
42.5	21.9	24.7	28.5	20.9	20.3
43.2	22.9	26.1	27.6	20.8	20.9
42.2	22.7	24.1	25.9	20.9	20.7
43.6	23.2	25.4	28.0	21.5	21.3
43.6	22.0	25.7	25.8	20.2	21.1
44.2	22.1	25.6	26.0	21.7	19.9
44.6	21.0	26.4	21.9	26.3	30.2

3. OUTLINE OF THE DEFORMED SHAPE GENERATION METHOD

The deformed shape generation method is based on free shape deformation, whose outline is as follows:

Step 1Obtaining the polygonal and skeleton information of a person with Kinect

Step 2Estimating the body parts that contain the respective polygonal information based on the skeleton information and allocating a deformation grid accordingly

Step 3Deforming the deformation grid with the designated parameters

Step 4Deforming the polygonal shape of the

person on the basis of the information on the deformation grid

Step 5Outputting the deformed polygonal information with a 3D printer

In step 1, the shape information of the self is obtained from two directions – front and rear –, and point cloud information is transformed into polygonal information with the use of Kinect Fusion[3]. And in doing so, other kinds of processing such as removing noise and filling up small holes are also carried out.

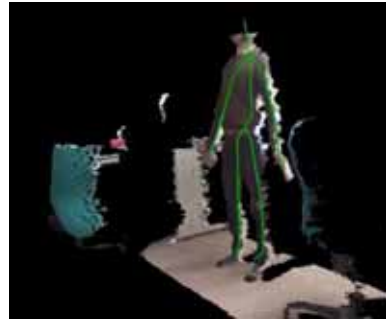


Figure 2:Example of information obtained in Step 2

Kinect has the function of estimating skeleton information, which is required for making character animation of obtained information on people. Step 2 allocates a deformation grid based on such skeleton information. Human bodies are roughly divided into six body parts (head, torso, left and right arms, and left and right legs).

The processing related polygonal shapes in steps 2 to 4 is carried out with the use of the 3DCG-making software Blender. Figure 3 shows an example of deformation. The image in the middle is the original shape, and the left and right images are the deformed shapes. The lines surrounding the human body on the left represent a deformation grid. In Figure 3. Image in the center is the original shape, and the left and right ones are deformed shapes.

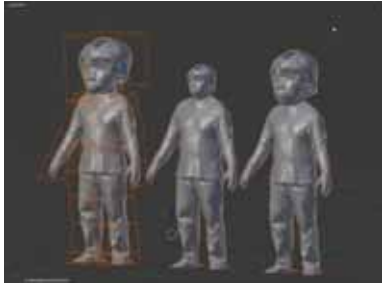


Figure 3: Deformed Figures and Original Shape

Step 5 transforms the polygonal shape into voxel data, and makes STL-format data on the basis of it.

4. RESULTS

Figure 4 and 5 show examples of figures made by applying the method of Chapter 2. The example of deformation (case 1) of Figure 4 does not involve large deformations.



Figure 4: Example of Deformation (case 1)

The example of deformation of Figure 5 was made by measuring the size of a deformed figure on sale and setting deformation parameters accordingly. In the example of deformation shown in Figure 5, the ratio of the sizes of each

body part assuming that the size of the head part is 1 is the following: torso: 0.51; right arm: 0.58; left arm: 0.60; right leg: 0.50; and left leg: 0.51. This ratio was calculated based on the average values of the deformed figures shown in Table 1.

Table 2: Parameters for Deformation in Figure 5.

head	torso	right arm	left arm	right leg	left leg
1.00	0.51	0.58	0.60	0.51	0.50



Figure 5: Example of Deformation (case 2)

Although Figure 4 and 5 do not show it very clearly, since the resolution of depth information obtained from Kinect is not very high, the faces have flattened shapes. Although many deformed figures on sale also have flattened faces, since they are colored, they do not look very unnatural. However, since the output examples shown above are not colored, it is difficult to identify the eyes in the face.

5. CONCLUSIONS

The paper outlined a method of obtaining information on human body shapes with Kinect and making deformed figures based on the obtained information.

Since simple free shape deformation is unable to process deformation of eyes in the face, it is needed to prepare deformations tailored to each part, by also making use of face image information and estimating the parts in the face.

Also, this method is highly dependent on the

functions of Kinect. Therefore, it is difficult to carry out processes that are difficult for Kinect. For example, since it is difficult to estimate the skeleton structure of a person wearing a skirt, it is difficult to apply the system developed this time to people wearing long skirts.

REFERENCES

- [1] M. Tonouchi et. Al. Otaku Vision. *NICO-GRAPH 2005*. 2005.
- [2] Yuto NARA, Genki KUNITOMI, Yukua KOIDE, Wataru FUJIMURA, Akihiko SHIRAI, Manga Generator: Immersive Posing Role Playing Game in Manga World, *Laval Virtual 2013* (2013).
- [3] Kinect Fusion ,
<http://msdn.microsoft.com/en-us/library/dn188670.aspx>.

ABOUT THE AUTHORS

1. Satoshi CHO is an adjunct instructor in Kanagawa Institute of Technology. His re-search interest is a computer animation. His email address is shisei@gmail.com.
2. Hisashi SATO is a professor in Kanagawa Institute of Technology. His research interest is computer graphics and interactive system. His email address is sato@ic.kanagawa-it.ac.jp
3. Hideki Yoshi graduated at Kanagawa Institute of Technology in 2014. Now he works in IT company as a programmer.

A METHOD STUDY OF RAPIDLY MAKING LARGE-SCALE TOPOGRAPHIC MAP

Aichen JIA¹ and Niannian WANG¹ and Pingping RUAN²

¹Faculty of Infrastructure Engineering, Dalian University of Technology, China ²Investigation and Design Institute, China Railway 11th Bureau Group CO., LTD, China

ABSTRACT: At present, it is relatively difficult to gain the basal data of large-scale topography rapidly, since conventional methods of making large-scale topographic map are not sufficiently smart and convenient. With the rapid development of spatial information and computer technique, researchers pay close attention to high efficient remote sensing data and intelligent computer. It is very urgent for us to study an intelligent and efficient method of making large-scale topographic map. This paper is based on the ecological recovery planning project of Heidingzi River, Changchun, Jilin. Since massive terrain data is lacking in this project, a method of rapidly making large-scale map by using data from Google Earth and process automation is presented. This method can quickly and efficiently make large-scale topographic map whilst its precision can satisfy the requirement of river ecological restoration planning. Firstly, elevation points are extracted from Google Earth. API Interface provided by Google Earth can be used to compile a program, with which we can gain three-dimensional coordinates of the terrain through calling its built-in function. Therefore the contour is obtained through interpolating and linking elevation points; then vegetation part is extracted from Google Earth image by utilizing the spectral characteristic of vegetation, followed by eliminating shadow of buildings using threshold segmentation; finally, the segmented roads and buildings are attained based on object-oriented classification, after which a complete map is made by combining contours and ground objects. After the completion of the map, images extracted from Google Earth can be added into the map to check its accuracy. The result shows that the map perfectly matches the original topography and features. This method has been implemented in Heidingzi River and it provides the terrain data for the ecological restoration project.

The Google Earth image of Heidingzi River is used in this example, shown in Fig1. With this method, we can make the complete map with combination of contours and ground objects, shown in Fig2. The example shows the method generated map rapidly without ground survey data and appropriate results were obtained.

Keywords: Topography, Contour Generation, Vegetation Extraction, Image Segmentation, Road Extraction, Building Extraction.

1. INSTALLATION

Topographic map is a graphic form which uses different landform symbols to describe the surface world. As the tool for people to understand the area of geographical knowledge, topographic map is also used in all aspects of the social production and living, such as urban

planning, analysis of urban land, the construction of the project and the flood risk analysis. Because the cost of topographic map mapping and making is very high and it needs a certain time period, and the area of the project topographic map is often very large, so in a timely manner to obtain the topographic map is very difficult. In this paper, we present a method of

rapid extracting topographic map which is based on Google earth and software automation information. This method can quickly obtain the topographic map data of the study area. Google Earth (GE) is a piece of fictitious tellurion software developed by Google. GE has integrated a variety of aerial photographs and the high resolution satellite image information. The highest resolution has reached 0.61 m (Quick bird). For big cities and building areas GE also provide the images, of which the resolution is about 1 m to 0.5 m. At the same time, this software includes the surface elevation information. GE provides the necessary basic data for the production of topographic map. In this study, GE remote sensing image of Heidingzi River ecological restoration is adopted as the image of study area, shown in Figure 1.



Figure1: Image of the study area.

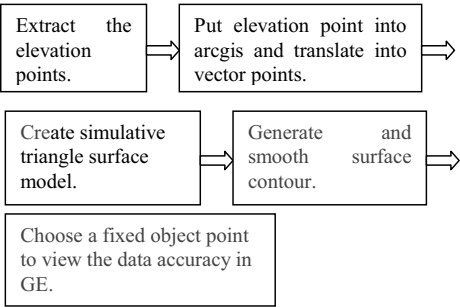


Figure2: Flow chart of contour generation.

2. MAKE THE CONTOUR

Contour is a closed curve which is produced by

connecting the points with the same height. It is the most basic and important terrain data of a topographic map. In this paper, the elevation extraction method is based on Google earth^[1]. Google earth has provided users with API interface. The API can be used for users to compile a program, with which we can gain three-dimensional coordinates of the terrain through calling its built-in function. Then we can imitate the 3D terrain surface through interpolating elevation points. Finally, the points on the surface with the same height are connected to generate the contour. All data is acquired and processed by the self-compiled program and Arcgis software automatically. Therefore this method improves the production efficiency of the contour. Figure 2 shows the specific operation flow chart. The generated contour map shows in figure 3.

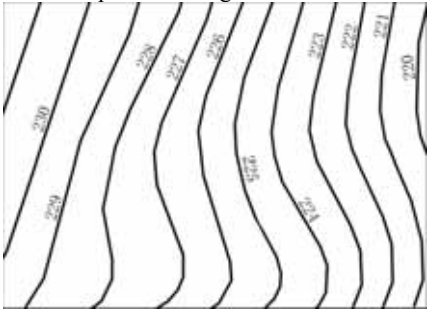


Figure3: Contours of the study area.

3. EXTRACT SURFACE FEATURES

There are a lot of methods proposed for the extraction of building features. For example, 1) Object-oriented classification method: The segmentation algorithm based on object is used to separate various features^[2]. Then according to the properties of different terrain features objects can be classified. 2) The method based on contour extraction: Through edge detection, edge connection, series of images analysis and processing, good result can be achieved^[3]. 3) The method based on shadow extraction: Firstly, the building shadow boundary is extracted for tracking the building seed points. Then we

can obtain the building of rectangular area by region growing^[4]. Using the above method to extract building good effect can be achieved even if there are certain restrictions. Since the remote sensing images only have red, green and blue three bands and the resolution is low, when using the object-oriented classification method, if the segmentation scale is small, the same building roof can be divided into petty parts by different spectral. However when the segmentation scale is larger, due to excessive merger, different kinds of objects can be mixed together. Contour extraction method is suitable for the building roof which has to have the same spectra and high resolution. The assisted shadow extraction is more suitable for the single building situation. Because of buildings gathering, a large number of buildings shadow stack together so that a satisfactory extracting result cannot be obtained. According to the analysis above, previous methods cannot apply to the current image data processing very well. Features of remote sensing images are mainly composed of four parts, namely the plant, shadow, roads and buildings. Since the difficulty of other three features extraction is less than that of the buildings, this paper adopts the method of firstly taking steps to extract the plant, shadow and road according to the characteristics of features, then extracting the building.

3.1 Extract the plant

The method of traditional plant extraction is to calculate the normalized vegetation index (NDVI), i.e. the ratio of the infrared band to the red band. As remote sensing images do not have infrared wavelengths, this method is not feasible. Comparing the three bands spectral features of plants and the others, the connection line of red, green and blue wavelengths mainly form a triangle. The green band spectral value is greater, and the red and blue bands spectrum values are relatively lower, as shown in figure 4. But the spectral values of the ground objects such as buildings and the shadow do not have

the obvious characteristic, as shown in figure 5. This paper proposes a new plant extraction method for infrared image, called multiple band ratio method based on pixels. The principle of the method is based on band operation. First, we use the band ratio of the green and red band to generate a ratio image, named picture 1. Second, we use the band ratio of the green and blue band to generate a ratio image, named picture 2. Third, we set the threshold to let the values of the picture 1 and 2 larger than 1.05. The intersection of the picture 1 and 2 after threshold segmentation is the plant area. Fourth, using morphological processing, we can fuse the outliers and other pixels around the outliers. The smaller point groups should be rejected in Arcgis. Finally, the complete plant mask is obtained. Figure 6 is the remote sensing image after removing the plant.

3.2 Extract the shadow and the sunny slope of the building

Since the backlight surfaces of the buildings do not receive the sunlight, their shadow cast on the ground is presented in deep black on the remote sensing images. The images of buildings in the southern slope are white in the remote sensing images. The buildings are mainly divided into three parts, the sunny slope, roof and shadows. After removing the sunny slope and the shadow, the roof of the building is separated. Some part of road connected with the buildings can be removed in the next step. The identifiable degree of the shadow and the sunny slope of buildings is high since their spectral signatures are different from that of the other objects. Therefore the method of gray-level threshold is used to separate them. Then by using morphological operation and the area calculation, small spots of the image can be removed to achieve a satisfactory result. Figure 7 is the image after removing the plant, shadow and the sunny slope of the building.

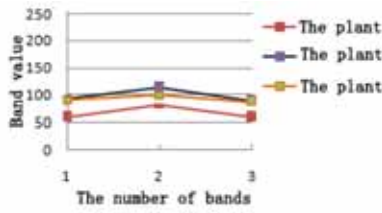


Figure4: Band value of vegetation.

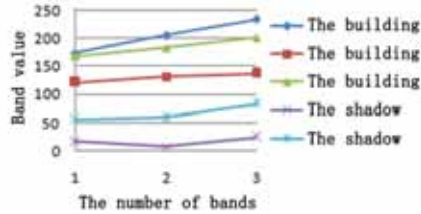


Figure5: Band value of other ground objects.



Figure6: Image without vegetation.

3.3 Extract the road

The shape of roads is relatively neat in the city. The spectral signature in one road is similar. Firstly, we can use the method of object-oriented segmentation on the image only including buildings and roads. Then we can set up different rules to extract different roads. This paper adopts the method based on shape and spectral heterogeneity. Segmentation algorithm adopted the data structure in the literature [6], combining with the merger guidelines in literature [7] to calculate the merging cost. Although the merging cost in literature [6] has considered the influence of shape factor, it

cannot adjust the shape and the proportion of spectrum according to the difference of the image. Therefore, it is more universal that we separately calculate the spectrum and heterogeneity in the shape as the cost of merger. The effect of image segmentation is better for the low resolution images. But for the rural road in this paper, since the resolution of the road is too low, the extraction effect is not very good. Roads can be manually added into the map. The calculation of merging cost is shown in the following formula.

Formula (1) calculates the total merging cost of the spectra and shape:

$$t = w \times h_{\text{color}} + (1 - w) \times t_{\text{shape}} \quad (1)$$

Where t is the total merging cost, w is the weights of the spectrum, h_{color} is the value of spectrum and t_{shape} is the merging cost of the shape.

Formula (2) calculates the shape merging price of the firmness and smoothness:

$$t_{\text{shape}} = w_{\text{com}} \times h_{\text{com}} + (1 - w_{\text{com}}) \times h_{\text{smh}} \quad (2)$$

Where w_{com} is the firmness of weight calculation, h_{com} is the difference of the firmness and h_{smh} is the difference of the smoothness.

Formula (3) calculates the difference:

$$h_{\text{diff}} = n_{\text{merge}} h_{\text{merge}} - (n_1 h_1 + n_2 h_2) \quad (3)$$

Where h_{diff} is the calculated difference, n_{merge} is the number of pixels after merging, h_{merge} is the property values of the objects after merging, n_1 and n_2 are respectively the number of pixels before merging, and h_1 and h_2 are respectively the property values of the two objects after merging. When calculating h_{color} , h is the variance (δ) and h_{diff} of each band can be weighted. When calculating h_{com} and h_{smh} , we can value $h = 1/\sqrt{n}$ and $h = l/b$. Where l is the actual circumference of the object, n is the number of pixels and b is the circumference of the object's external rectangle.



Figure7: Image of buildings.

3.4 Extract the building and compound the topographic map

Using the extracted roads as a mask, the road can be removed in the image. Therefore the image only contains the buildings. It can be converted into binary image by threshold value method. Then the binary map is converted into vector graphics in Arcgis. Using the building simplification tool provided in Arcgis, the sawtooth on the boundary of the extracted buildings can be repaired to get a regular building shape. Finally the complete topographic map can be obtained by adding the extracted feature into the complete contour, shown in Fig 8.

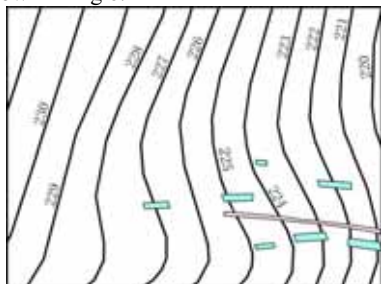


Figure8: Map of the study area.

4. CONCLUSIONS

In this paper, various image processing techniques are comprehensively used to make contour and extract various objects so that a fast and convenient method is provided to make topographic map and achieve good effect. Compared with other methods, this method has greatly improved the efficiency, but it still need

to be improved in precision and detail. For example, the first, the contour accuracy entirely depends on the elevation data of Google Earth, which cannot be improved by users. The second, the building can be slightly dislocated during the process of simplifying the sawtooth or radian on the edge into regular shapes. The third, the extraction result of the rural road is not very good. In the next step of research, we should focus on how to improve data accuracy and edge simplified in order to produce topographic map with better quality.

REFERENCES

- [1] Jiang K, Gong X P. Development Details of Google API: Combination of Google Maps and Google Earth [M]. Beijing: Electronics Industry Press, 2010.
- [2] Xu C R, Ge S Y. Object-Oriented Remote Sensing Image of Building Extraction [J]. Urban Geotechnical Investigation & Surveying, 2011, 84-86.
- [3] Wang Y G, Wang Z Y, Wang H B. High resolution remote sensing image based information extraction of building outline [J].
- [4] Yu D F. Research on Shadow-based Automatic Extraction of Buildings from Remote Sensing Images [D]. National University of Defense Technology, 2008.
- [5] Zhang Y J. Image Segmentation [M]. Beijing: Science Press, 2001.
- [6] Robinson D J, Redding N J, Crisp D J. Implementation of a Fast Algorithm for Segmenting S-AR Imagery [R]. http://www.dsto.defence.gov.au/c-orportate/reports/DSTO_TR_1242.pdf, 2007.
- [7] Baatz M, Schape A. Multiresolution Segmentation: an optimization approach for high quality multi-scale image segmentation [EB]. <http://www.agit.at/papers/2000/baatz-FP-12.pdf>. 2000.

ABOUT THE AUTHORS

1. Aichen JIA (1962~), born in Liaoning, China, PhD, Professor, Research Direction: Engineering Graphics, Water environment and Ecological Management.
2. Niannian Wang(1989~), born in Henan, China, Master, Research Direction: Engineering Graphics and Water environment.

THE METHODOLOGICAL BASES OF GEOMETRIC MODELING OF MALTIFACTORIAL PROCESSES

Vladimir VOLKOV¹, Olga ILYASOVA¹ and Margarita CHIZHIK²

¹Siberian State Automobile and Highway Academy, Russia

²Omsk State Institute of Service, Russia

ABSTRACT: Currently, modern manufacturing techniques allow you to create different material systems (MS) for clothes with a fundamentally new operating properties that best meet the inquiries of the consumer. A special group should select a system with volume materials, which are used in the manufacture of many types of clothing due to the wide range of technical parameters and high operating characteristics.

Keywords: an optimization, technological processes, a multidimensional space, Radischev's drawing.

Practice shows that in the process of manufacturing of garments give rise to numerous challenges posed by the multi-levels and multicomponent. Examples of such tasks include: the task of forming properties and characteristics of investigated object. the selection of optimal treatment of materials which provide the specified level of indicators of the quality of the finished product; objective quality assessment and prediction of properties and characteristics of an object depending on the parameters of forming and operating conditions. At the design stage of their decision by traditional methods, is complicated, and in some cases becomes virtually impossible, particularly because of the need to identify patterns associated with taking into account many factors and establish quantitative relationships that define the metrics property of products taking into account the conditions of operation and using.

For solving the task of this class are used extensively by classical methods of mathematical modeling. However, the models are not practical, because most of the focus on research objectives are characterized by a large number of mathematical operations and the lack of a Visual representation of the object.

The lack of methods to solve the above tasks and information processing tools, adapted to the realities of production does not allow professionals to fully make the best technology decisions, provide production with specified performance properties, dematerialization, time and increasing productivity.

It follows that research aimed at the development of methods, models and tools for computer-aided design of multicomponent systems of MSM for textile industry goods are relevant.

The aim of this work is to develop the theoretical basis and practical methods for modeling multifactor, multiobjective systems and processes that ensure high production efficiency and the quality of the finished product.

In the process of making modern clothing packages from source materials are converted into new, complex objects, distinguished according to the way they connect to and retrieve, as well as a constructive solution. Therefore, from the perspective of systematic approach package should be considered as a complex, multi-component system consisting of several layers of materials (elements), each of which has a strictly defined properties and characteristics. Relationship between the materials is provided in various ways of connection (thread, glue, welding, etc.). When a certain combina-

tion of elements of a particular type of connection tells the system a set of new, non-original material properties.

Thus, in this paper, under the system of materials (SM) will understand the package, interconnected in various ways. Multi-component system materials should be considered as containing no less than three different in appearance, composition and structure components. The number of components depends on the type, purpose and operation of the sewing machine.

This object belongs to a class of complex systems and can be represented as multi-parameter model (Fig. 1). Elements of the system are different materials $x = \{x_1, \dots, x_k, \dots, x_n\}$ with a variety of quantitative and qualitative characteristics $p^{(k)} = \{p_1^{(k)}, \dots, p_{m_i}^{(k)}\}$, where $p_i^{(k)}$ - the i -th characteristic of the k -th element of the system, $k = \overline{1, n}$. The control information $b = \{b_1, b_2, \dots, b_l\}$ is contained in various kinds of specifications and technical documentation. Information about the designation of manufactured products, operating conditions $c = \{c_1, c_2, \dots, c_f\}$ and modes of operation of technological equipment $d = \{d_1, d_2, \dots, d_r\}$ is used as a tool of the system. MSM with specified indicators of properties are the output parameters $y = \{y_1, y_2, \dots, y_q\}$.

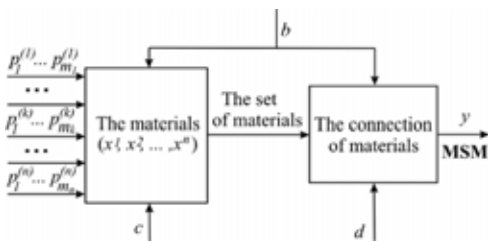


Figure 1: The model of object of study.

As noted above, the complex new system gets the properties of the connection input thread, glue, welding and mechanical means. To date, in the production of garments is the

primary way connection is wound. This is due to the versatility of its use, the variety of parameters and properties of the comparative ease of manufacture, a wide selection of equipment for its receipt, etc.

Ongoing maintenance and expansion of assortment of textiles and sewing thread requires the selection of optimum modes of connection of details, providing joints with the specified quality characteristics. Ensuring a certain level of quality is a complicated multi-factor connection task associated primarily with the search of optimal parameters of their education.

The choice of optimal connection of clothing details on a large number of works. Their analysis showed that for solution of optimization tasks processes compounds widely used mathematical methods of experimental design, linear programming and regression analysis. Authors of the mathematical model in the form of the equations of regression based on optimization criteria from the managed connection settings. However, even if these equations are adequate, in order to find the best possible process parameters requires a decision by a single task, for example, set the grid of possible parameter values for each site, and compare the values obtained with the specified quality. Solution to this problem is possible only with the use of computers, the current software does not provide a graphical visualization of parameters $k > 2$. In addition, mathematical models are rated for each optimization individually, and it requires considerable expenses in performing calculations.

All this makes the topical question of the elaboration of the methods of the build and analyze models that take into account the many independent parameters, and provide visibility to the experimental data.

Statement of the problem. Let the system components are the materials and processes, each of which has a set of properties and characteristics. Select the necessary parameters of $\{x_1, x_2, \dots, x_n\}$ depend on quality criteria of $\{y_1, y_2, \dots, y_m\}$. Based on the results of the

measurements or observations for each set of parameters in tabular form of quantitative values of quality criteria [1].

Each criterion of quality, you can set the desired value y_i^* and the possible values of the deflection δ_i from this value y_i^* :

$$|y_i - y_i^*| \leq \Delta_i \quad (1)$$

Required for a given set of values of the indicators (benchmarks) $\{y_1, y_2^*, \dots, y_m^*\}$ find valid scope parameters in multidimensional space.

Depending on the number of parameters and criteria for the quality of the task can be the following: -the number of parameters is greater than n – number of criteria;

-the number of parameters is equal to the number of criteria;

-the number of parameters is less than the number of criteria.

For each of these cases can be obtained by different geometrical models.

As an example, consider the application task of sewing production associated with finding acceptable options area of the corresponding optimal regimes thread compound textile material.

There is a system, the elements of which are textile materials with a certain set of features and technological process of processing, for example, process connection thread Outline of dynamic parameters of connection: x_1 - thread-stitch length; x_2 - the tension of the needle thread; x_3 - thick sewing threads. The degree of influence of process parameters on the quality criteria may be different, so you should be the most important. Parameters such as operating force paws on the fabric, the diameter of the needle made permanent, ensuring high-quality stitch.

Quality criteria are indicators characterizing the mechanical properties of connections: y_1 - breaking load thread suture in the transverse direction; y_2 - rigidity 3thread seam; y_3 - breaking load thread seam in the longitudinal direction.

Set the quality indicators for the following

values: $y_1 = y_3 = 200$; $y_2 = 170000$.

Required for a given set of values $\{y_1^*, y_2^*, y_3^*\}$ find valid scope parameters, modes in multidimensional space.

Based on the results of the experimental data presented in tabular form, for each set of connection settings are defined, the values of the indicators selected as optimization criteria $y_i(x_1, x_2, x_3)$.

If one quality criterion, the task one-criterion. For its solution is the use of regression analysis.

Task considerably more difficult when you want to simultaneously record multiple criteria optimization [2].

Geometrical multiple-criteria can be summarized as follows: set the grid of possible parameter values, then each node (x_1, x_2, x_3) this grid is experimentally found value y_i . Construction of Lagrange interpolation function, passing through the point of y_i , find the equation of the surface in a multidimensional space. The specified values y_i^* and Δ_i can determine the allowable area of D_i . The crossing will be the total acceptable D_i area parameters simultaneously by multiple criteria. Solution to this problem is possible only with the use of computers, the current software does not provide a graphical visualization to number more than two parameters.

Ensuring clarity can be achieved using multivariate descriptive geometry, introducing the source data and the results obtained in the form of a graphical model that is linear or non-linear subspace of the multidimensional space and the relations between them. In descriptive geometry of multidimensional space scientists suggested a number of ways and justified the drawing of multidimensional objects based on projection apparatus. Your choice of format is mainly determined by the goals and objectives of the study.

Analysis of the literature in this area has shown that for solutions to complex technological tasks are used models in the works. However, the most convenient for a graphic representation of the model of the multidimensional

space is Radischev's drawing at which point space is defined by (n-1) projections (Fig. 2). A point in space E_n orthogonally project the (n-1)-2 planes of coordinate system, where x y z t of sharing a common axis, such as h. Received orthogonal projection into 2-plane drawing that matches with a 2-axis planes, so that the (n-1) perpendicular axes x, appear in the same line and the x-axis is displayed as perpendicular to the (n-1) direct X_i ($i = 1, 2, \dots, n - 1$). Everything said above is true when x is selected instead of any other axis.

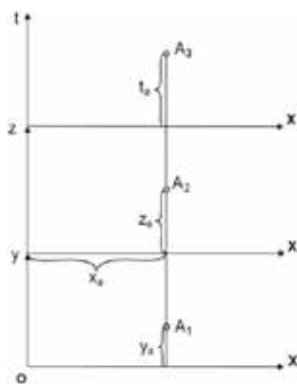


Fig. 2 - Model four-dimensional Radyshchev's space

This model is most suitable for simulation and optimization of a various degree of complexity. Improvement, development and application of multifactorial modelling Radischev's drawing processes is devoted to a number of works, among which special attention should be paid to study V. N. Pervikova, N. F. Chetveruchin, V. Y. Volkov, V. U. Yurkov and others. Their analysis showed that this model has been successfully applied to modelling multifactor, multicomponent systems in physics and chemical analysis, the study of the properties of such systems use methods of the enumerative geometry and the theory of parameterization [3].

The main idea of the work is to create a

formalized mathematical (geometry) apparatus for solving complex problems in analysis and synthesis of data and construction of algebraic manifolds, which with a sufficient degree of approximation can be used for geometric modeling of various processes into physical-chemical analysis of multicomponent systems.

However, despite the development in this direction, remains a topical issue on the validity of the objectives of the models, there are no methods of solving multifactorial, multiobjective tasks, and software to automate the build of geometric models by taking into account the requirements of the applications.

From the existing literature, it is known that graphical models of multidimensional spaces, mostly are built similar to design graphical models of 3D space. This increases the dimension of either m-plane projection, or projection center. Increase in the size of objects does not allow visibility in geometric studies.

Based on the analysis of building graphical models spaces of varying size, the structural characteristics and class of tasks, which are solved on these models, professor V. Y. Volkov formulated a system of axioms, which includes elements of the theory of parameterization and methods of multidimensional enumerative geometry.

Axiom 1. The dimension of the space of main objects of types of manifolds and image space is identical with the core objects are equal.

Axiom 2. Structural characteristics of varieties of the main objects of types and fundamental space objects space images are equal.

Axiom 3. Model of space images is defined relative to the geometric problems which are solved in the prototype.

This axiomatics allows using a formalized apparatus, test the constructed model.

Let's say that simulated space whose main object is the E^4 0-plane (point) and the formula of Grassmann, is a four-parameter (2):

$$D_4^0 = (0 + 1)(4 - 0) = 4 \quad (2)$$

This space is a prototype. For the space of images, choose the top three 0-planes, which lie on the beam. To select the beam to beam you need to ask one parameter, and three 0-plane will be determined by three parameters. Then diversity triples 0-planes, on the beam, the beam will be a four-parametric. It follows that the proposed model meets the first axiom.

If you apply the apparatus enumerative geometry, it can be seen that their structural characteristics are equal to one, then there is space images and the prototype will be linear. Thus, the second axiom is satisfied.

Finally, a prototype is defined relative to the positional problems. Then you should choose the simplest model on which tasks the incidence and intersection geometry.

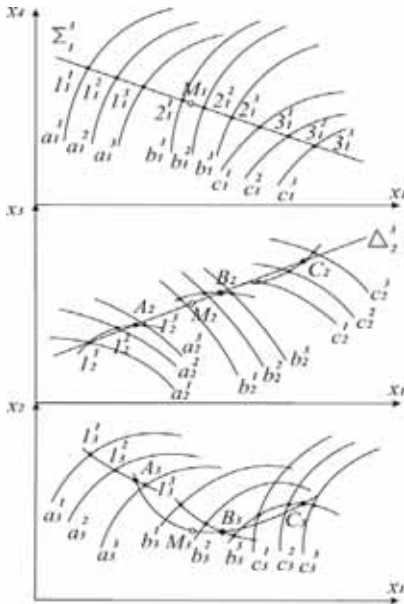


Fig. 3 - Hypersurface model of two parametric of manifold 1-surfaces

Axiomatic method for graphical models revealed the adequacy of model simulation space and develop a formal design algorithm for op-

timal models of spaces of varying dimensions, structural characteristics of geometric problem over the field of real, complex and hypercomplex numbers .

Proven in terms of the axiomatic theory of the adequacy of this drawing as a model of multidimensional space allows you to reliably use it to solve optimization problems.

For reasons of clarity of images and solve certain geometric problem on a drawing surface for the job, Radischeva use frames. This approach is applied when graphical model analyzed process is a hypersurface, which can be set with the discrete number of experimentally obtained 1-surface or using 1-surfaces that are built based on tab regression equation that defines the hypersurface [4].

Consider the particular cases set in the drawing, Radischev's hypersurfaces axiom. Set the hypersurface in the drawing (Fig. 3) two parametric multiplicity surfaces $a_1 a_2 a_3$; $b_1 b_2 b_3$; $c_1 c_2 c_3$.

Let's say that two projections set 0-plane (M1, M2). You must build the third projection is the assumption that 0-plane m belongs to the hypersurface. Suppose also that hypersurface continuous smooth and has no special features. Then as 1-surface can take 1-surface of the second order. This assumption can be taken in a small area of the process that is modeled by hypersurfaces. Then select the 2-surface that passes through 0-plane M1 and is projecting a hyperplane Σ_1^3 .

Let's build this 2-surface of the 2-plane projections $x_1 x_3$ and $x_1 x_2$. The latter will be presented to three 1- surfaces, $1^1 1^2 1^3$, $2^1 2^2 2^3$, $3^1 3^2 3^3$.

Next we build 1-surface ABC that also passes through the M2 and lies in projecting a hyperplane Δ_2^3 . Get her projection of $A_3 B_3 C_3$ on the projection 2-plane $x_1 x_2$. Because according to the problem conditions 0-plane M belongs to the hypersurface, its projection of the M3 will belong to the $A_3 B_3 C_3$.

The tab the regression equation is useful to

take some independent variables constant, discretely changing their values, and the rest will be determined from calculations. Then the hypersurface in the drawing will be 1-surface degree or level surfaces (fig. 4, 5).

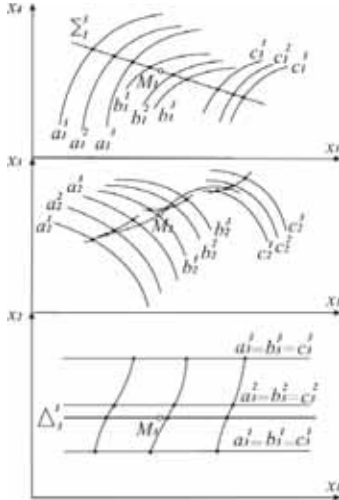


Fig.4 - Setting the two-parameter manifold hypersurface 1-level surfaces

Fig. 4 hypersurface given two parametric manifold 1 level surfaces, Fig. 5 hypersurface given two parametric manifold 1-surfaces of the double standard. In each of the drawings illustrates the 0-plane M which belongs to the hypersurface. On the basis of what has been accepted as a 1-surfaces choose smooth continuous surface 1, it follows that they can build as smooth contours or interpolate splines.

According to the work of the hypersurface can be considered as given in the MMP, if graphically is one of two tasks:

- check the identifiable 0-plane of the specified drawing hypersurface;
- to complete the missing projection 0-plane, if it belongs to the hypersurface.

Generalized the algorithm performed above builds can be formulated as follows:

1. The given 0-plane sign in $(n - 2)$ -surface, which is projecting a hyperplane passing through one of the specified projection 0-plane.

2. Build projection $(n - 2)$ -surface on the other views, the hypersurface.

3. These builds with a selection subsequently to the surface dimension $(n - 3)$, and $(n - 4)$ etc., underlying the project hypersurface, passing through the projection of this 0-plane, continue until you come to the surface 1.

4. When the conditions of the theorem (a) needs to check if the latest projection given 0-plane; in the case of (b) builds the latest projection 0-plane is selected on this 1-surface.

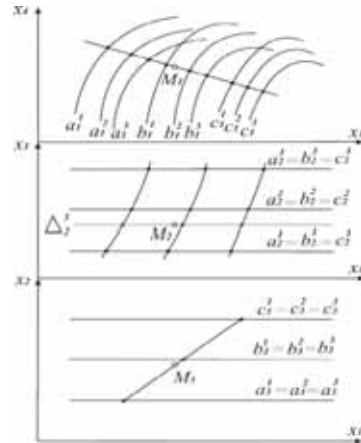


Fig.5 - Setting the two-parameter manifold hypersurface 1- surfaces double level

Consider the sequence of building common elements of two or more objects in a drawing (trench).

When you build a surface crossing $m1$, $m2$ -surfaces in n -dimensional space, the original $T1$, $T2$ -surfaces must meet the dimensional formula $T1 + T2 - w = k$, where $k \geq 0$. Therefore, the original data is checked for the specified condition.

Next, have the smallest size of T -surfaces in projecting the hypersurface, for example

$\tau_1 \geq \tau_2$ find $T_i = (n - 1 + 1 - p)$ -surface intersection. Performing the above operation to a finite number of times, you can reduce the dimension of the original T-surfaces so that $T_i = k$. To look at the drawing the intersection of two 2-surfaces abc (a1 b1 c1, a2 b2 c2, a3 b3 c3);def (d1 e1 f1, d2 e2 f2, d3 e3 f3).. Make 2-surface (a) (b) (c) projecting the hypersurface, which - 1 protection plane (fig. 6).

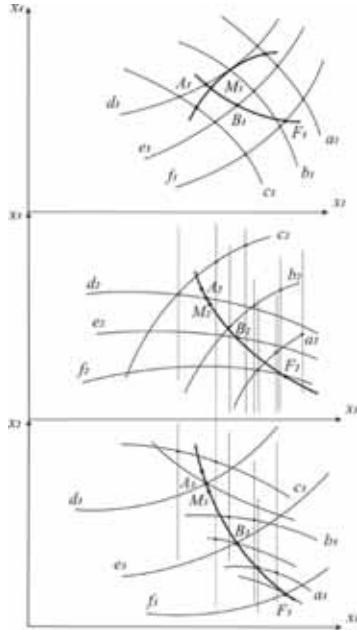


Fig. 6 - Construction of the intersection of two 2-surfaces in general position

Find 0-plane intersection (A1, A2, A3) hypersurface with 1-surface; 0-plane B (B1, B2, B3) crossing the hypersurface with 1-surface e; 0-plane F (F1, F2, F3) crossing the hypersurface 1-surface f (f1, f2, f3). 0-plane ABF establish the 1-surface owned by the 2-surface def, 2-surface ABC is a hypersurface. Define one of 0-planes M (M1, M2, M3) crossing 1-surface and ABF and 2-surface and abc.

The decision is easier if one of the

2-surfaces is a 2-plane (fig. 1), or both two 2-surfaces are 2-plane. For these cases, algorithm 0-plane intersection set 2-surface from the formulated above algorithm should build 2-plane intersection of two planes. In this paper verified that when applying the first and second algorithm received the same 0-plane.

The developed algorithms solve the positional problems constructively to allow the model to find common elements of two or more m-dimensional surfaces, that can be used when solving technical tasks of multi-component systems and processes.

Thus, the essence of the building optimization model in a Radischev's drawing is as follows. For example, suppose you have a multi-component system (x_1, x_2, \dots, x_n) and several independent optimizing factors of quality criteria (y_1, y_2, \dots, y_k), where y_i – settings and optimizing based on factors that have a set of tabular data resulting from measurements, observations, and so on to determine these parameters depending on the values of optimizing factors:

- set the hypersurface optimizing factors by selecting curves of a particular class, have a specific location relative to the starting points for each factor f_i ;
- Select and specify the optimal values of $f_i = f_{i \text{ OPTIMUM}}$ will be geometrically hypersurface level;
- determine the intersection hypersurface with level hypersurface in n-space and optimize settings for the specified area of the optimal values of factors.

Consider the effect of the above algorithm in solving applied tasks of sewing production-optimization process 3thread connection details.

The main objects of study were the thread connection on the machine stitch 1022-m class. The grinding was done at angles of 15° , 30° , 75° to the warp threads, because in these areas the stitches are the largest operational load. Samples of joints were fabricated using sewing thread of different linear density of tissues of

different fibre composition with polyurethane thread. Characteristic of sewing thread is represented in the table.

As noted above the process of formation of the seam was chosen: thread x 1-stitch length; x 2-the linear density of sewing thread; x 3-the tension of the needle thread. The reason for choosing these parameters is that they most often are selected and adjusted in the process. Operating force paws on the fabric and the diameter of the needle made permanent, ensuring high-quality stitch.

When conducting your experiment settings education thread seam ranged within the following limits: x 1 from 2 up to 4 from step 1; x 2-28, 32 and 37; x 3 from 0.2 to 0.6 in increments of 0.2.

Quality criteria were: U1-breaking load thread suture in the transverse direction; U2-rigidity thread seam; Y3-breaking load thread seam in the longitudinal direction. Determination of breaking load and stiffness thread seams is carried out by standard methods.

It should be noted that the number of technological parameters and optimization criteria may increase depending on the requirements of the applied tasks.

When designing a hypersurface is necessary in finding the type of curve or surface of a certain class and order, some located on the data points. This is done by calculating approximating or interpolation curves, equal all or several points from the original dataset. If the experimental data obtained with some accuracy using interpolation can give an approximation of bad quality. It is therefore advisable to build a close function to reduce and smooth the effects of measurement errors and the number of anchor points. This smoothing is achieved when constructing function approaching method of least squares.

In Figure 7 is an optimization model of linear shuttle stitching fabric connections made at an angle of 30° to the warp threads. Red line detected levels of optimal values of optimiza-

tion criteria (hyperplane). For example, the values of the optimal level of breaking loads taken equal to U1 Optim. = Y3 Optim. = 200 both lengthwise and crosswise, and rigidity – f2 Optim. = 170-103.

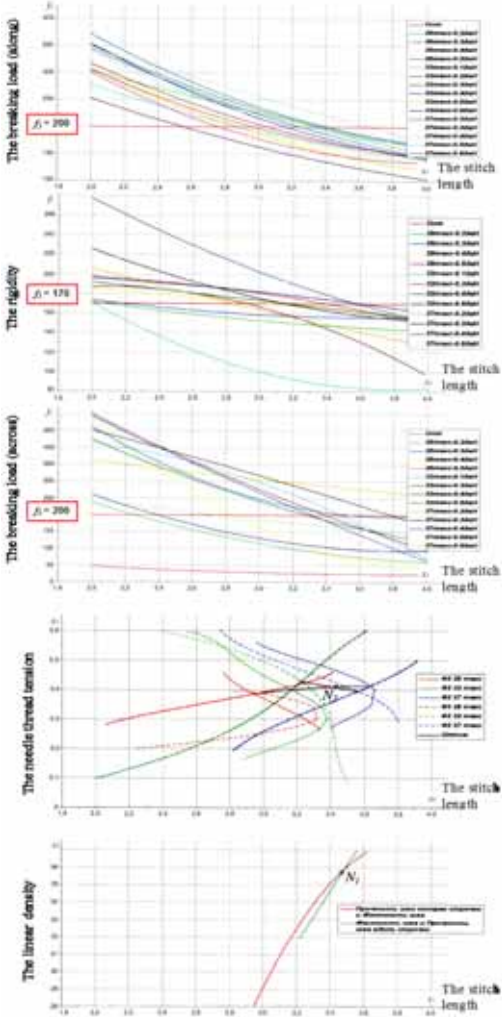


Fig. 7 – The model of multidimensional space - Radyshchev's drawing

As a result of crossing the hypersurface with hypersurface' \hat{u} level optimization area is the point N (N1, N2) whose coordinates $h1N = 3.5$; $h2N = 36$; $h3N = 0.4$, determine process parameter values education 3thread seam, whose f1 Optim. = 200, f2 Optim. = 170-103 and f3 Optim. = 200.

To automate the drawing of geometric patterns designed by a computer program that enables the optimization of multiple processes with a different number of parameters depending on the values of quality indicators.

Thus, drawing Radischev enables the connection to the process of visually 3thread as geometric models of many variables ($n \geq 3$) and allows you to determine the quality of the weld while you post multiple criteria optimization.

CONCLUSIONS

1. The algorithm for determining the area of intersection hyperplane with the hypersurface Radyshev's the drawing, which yields results making applications in the form of graphic optimization models is formulated.
2. Automation software and drawing optimization models multifactor processes is created.
3. The graphics optimization model of the process thread connection parts of clothing, which allows to vary the values of the basic parameters and filter modes that provide desired properties of the weld is built.

REFERENCES

- [1] Volkov V.Ya. Graphic optimization models of multivariate processes: a monograph / V.Ya. Volkov, M.A. Chizhik. - Omsk Univ OmGIS, 2009. - 101 p.
- [2] Volkov V. Ya. Multivariate enumerative geometry: a monograph / V.Ya. Volkov, V. Yu Yurkov. - Omsk Univ OmPU, 2008. - 244 p.
- [3] Volkov V.Ya. The theory of the parameterization and modeling of geometric objects

of multidimensional spaces and its applications: Abstract. dis. Dr. of technical. Science / V. Ya Volkov. - Moscow Aviation Institute, 1983. - 27 p.

- [4] The course of descriptive geometry on the base of geometric modeling: a textbook / V. Ya Volkov, V. Yu Yurkov, K. L. Panchuk, N.V. Kaygorodtseva. - Omsk Univ SibADI, 2010. - 253 p.

ABOUT THE AUTHORS

1. Vladimir Volkov - Dr.eng.sc., professor. Head of Dep. of Descriptive geometry and Computer Graphics Federal State Budget Educational Institution of Higher Vocational Training "The Siberian Automobile and Highway Academy (SibADI)". e-mail: volkov_vy39@mail.ru postal address: Dep. of Descriptive geometry and Computer Graphics Federal State Budget Educational Institution of Higher Vocational Training "The Siberian Automobile and Highway Academy (SibADI)", RF, Omsk, 644008, pr. Mira 5.

2. Margarita Chizhik – PhD, professor, professor of Dep. of Designing of garments of Federal State Budget Educational Institution of "Omsk State Institute of Service". e-mail: margarita-chizhik@rambler.ru postal address: Dep. of Designing of garments of Federal State Budget Educational Institution of Higher Vocational Training Omsk State Institute of Service, RF, Omsk, 644043, st. Pevcova 13.

3. Olga Ilyasova - PhD, associate professor of Dep. of Descriptive geometry and Computer Graphics Federal State Budget Educational Institution of Higher Vocational Training "The Siberian Automobile and Highway Academy (SibADI)". e-mail: ilyasovaolga@mail.ru postal address: Dep. of Descriptive geometry and Computer Graphics Federal State Budget Educational Institution of Higher Vocational Training "The Siberian Automobile and Highway Academy (SibADI)".

way Academy (SibADI)", RF, Omsk, 644008,
pr. Mira 5.

MODELING OF SHIP HULLS WITH NURBS CURVES AND SURFACES

Francisco PEREZ

Technical University of Madrid UPM, Spain.

ABSTRACT: This paper presents a mathematical method to define a 3D computer ship hull shape from its 2D Lines drawing with the use of NURBS curves and surfaces, so the procedure of the computer reconstruction of the hull shape can be automatized. Reverse engineering of ship hulls is requested by the lack of old manufacturing documentation, the need to develop CAD models based on hand crafted physical models or the wish to produce a new ship hull reusing the shape or part of the shape features of an existing design, or just only its shape coefficients. This research works with a discrete number of points or offsets, obtained semi-manually from a lines drawing, or from rastered data points obtained from an image of the drawing. The paper explains how to reconstruct the 2D lines of the hull with 3D NURBS curves, and how to use B-splines surfaces that fit these curves. NURBS surfaces can be easily exported to commercial and proprietary CAD, and are a standard environment for engineering and architecture applications. Techniques developed in this paper are useful for engineers, architects, students and small ship modelers that want to fit surfaces through a set of space ordered data points, for the definition of a 3D CAD Model. Actual CAD programs do not produce this kind of fitting automatically.

Keywords: NURBS, Reverse Engineering, Curve fitting.

1. INTRODUCTION

A research area in computational geometry is the development of algorithms that can reconstruct the 3D surface of an object represented by its 2D lines drawing definition. Different methods exist focused on objects with planar faces. In this work, a method for the reconstruction of objects that have not only planar faces but also curved surfaces is presented. The method is focused on ship hulls, that present typical curved faired surfaces, considering their 2D typical definition based on a lines drawing that contains three different families of planar cross sections on the hull (Fig. 1). The definition of a ship hull is one of the most restraining processes in initial ship design.

Single 2D lines drawings provide a "simple" way of illustrating 3D objects. In the case of a curved surface like a ship hull, three different

families of parallel cross sections are used: stations (left of Fig. 1), waterlines (bottom of Fig. 1) and longitudinal sections in profile view.

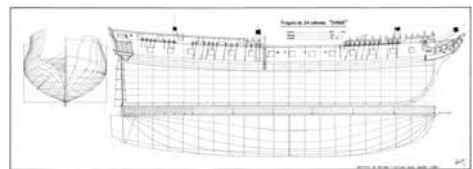


Fig. 1 Example of a lines drawing

The final result of the modeling will be the modeling of the ship hull with a surface or set of surfaces, that represent well the mentioned curves. Applying the method described in this paper, the 3D ship hull can be modeled as depicted in Fig. 2, where the ship corresponding to the lines plan of Fig. 1, has been modelled with NURBS surfaces applying

the method.

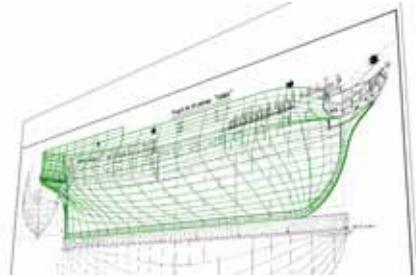


Fig. 2 Modeled ship hull

There are several papers about 3D reconstruction from single 2D lines drawings, not focused in ship design, i.e. [1], [2] that are based on the reconstruction of objects formed with planar faces. Little work is available to manage with curved objects, i.e. [3], [4]. In [3], different 3D patterns were used as inputs for the 3D object retrieval. The method needs the user to manually add different curves on the object surfaces, and work with the reconstruction of objects with low curvature. The method presented in [4] can work with simple curved objects, but is focused on architectural modeling where surfaces are planar most of the times.

In the method presented in this paper, the offset data points are obtained from the ship lines drawing, which can be made in different ways. Once that the data points are obtained, the method continues with a constrained B-spline fitting of the points obtained preserving the tangent angles at their ends. This fitting uses a parameterization based on the distance from the data points to the B-spline. This is very useful for manufacturing purposes as it enables an assessment of the maximum or average deviations from the original analytic curves. A lofting surface from the previous B-spline curves defines the final hull surface. Single surface hull representations are best suited to small craft forms like those covered in this paper because their hull shape is more continuous. The single surface representation demonstrates that the method presented here is

reliable and robust.

During the final stages of the method, the designer has the flexibility to control the number of control points on the surface and to enable a later gradual refinement of an initially simple surface to form a more complex hull.

2. OBTAINING THE DATA POINTS

There are up to three steps in the method of modeling an existing ship hull based on its lines drawing, which are common to any reverse engineering process. The first step is to use some input device or technique to collect the raw geometry of the ship hull. This data is usually in the form of (x, y, z) points on the object relative to some local coordinate system and are called offsets. These points may or may not be in any particular order, but the connection sequence and the density of offsets are very important in order to obtain good results.

The second step is to use a computer program to read this raw point data and to convert it into a usable form. This step is not as easy as it might seem. The third step is to transfer the results into some 3D modeling process, in this case using NURBS surfaces.

For this particular problem of ship hull modeling, there are two methods to obtain the raw data points:

2D Digitizer Methods: These technique consist of a flat, tablet-like part that hooks up to your computer. Once you tape your drawing or picture on the flat tablet, you use one of many types of connected input pointing devices (pen, puck, or stylus) to trace the geometry you want into the computer. You may use a program that comes with the tablet or you may use a general-purpose 2D or 3D graphics design program.

To input the geometry, most programs will have you position the pointing device at closely spaced positions along each line or curve in the drawing and input the 2D (x, y) point by clicking a button on the pointing device. A pen input device is often used if accuracy is not

critical or if you have a lot of points to enter. A “puck” type of pointing device with very fine crosshairs is used for very accurate work. A tablet is good for inputting lines and curves into the computer. All tablets also allow a stream mode where (x, y) points are continually sent to the computer as you move the stylus. In this case, ordered sets of points are obtained, following the order that the user digitize. The points correspond to the stations of the hull.

2D Scanners: These common devices work like digital photocopiers and are good for small drawings or pictures. They are fast, but they only get the drawing or picture into the computer as a matrix of color dots (a raster or bitmap image), just like on the computer screen. The resolution might be very high, but the raster format of the geometry may not be in a useful format. If a drawing consists of a number of lines and curves that you want to work on or use in some kind of 2D or 3D geometry modeling program, then an additional software is needed to convert the raster image into some kind of line or “vector” format.

There are two ways to do this. One way is to use a raster to vector conversion program. These programs look at the raster image and try to connect the dots to form lines or curves that can be transferred to your design program. As you can imagine, these raster to vector conversion programs can get easily confused if many lines or curves cross each other on the drawing. After this conversion, you might have to spend a lot of time in your design program cleaning up the mess. It might be faster to use a 2D digitizer tablet to input the data. Another way to convert the raster data to vector data is to use a design program that can read the raster data and display the picture as a background image. Then you can use your design program to recreate the vector geometry by “tracing” over the raster image the different stations. This is kind of like doing the digitizing right on the computer screen, enables to obtain an ordered connection between data points, but depend on user experience on selecting points

and can be quite tedious.

Knowing the source and quality of the offsets to be matched is very important, as well as having access to the original source to check offsets and to get more data.

Time spent preparing the offsets to be used in the modeling process is time well spent. Two key points are that the data should be error-free, and that the vessel should be well defined by the offsets. For the user’s eye may be able to easily interpolate the shape of the hull from a small number of offsets, but the computer will not be able to do so.

3. APPROXIMATION OF THE OFFSETS

The process of automatic or semi-automatic fitting of a surface to offsets requires some sort of mapping of each offset point to a corresponding point on the surface. One approach that has been used is to first fit a spline to stations, thus mapping the offset points to points on the splines, and then to pass longitudinal splines through these station-wise splines to define the hull. Often this method uses interpolating splines in either the station-wise, longitudinal, or both directions (interpolating splines are splines in which the curve passes through the control vertices). In this case, each offset point on a station would be a control point, to guarantee that the curve would pass through it.

The problem arises in the oscillatory nature of these splines, which means that although the curve will pass through the offset points, it may oscillate between them. Even if the stations are fit with approximating splines, the same problem can arise in the longitudinal fairing, which may contain oscillations. Also, this process tends to fair across any features such as chines or knuckles, losing these features in the process. In summary, if interpolating splines are used, and control points are placed on the offset data points, the match to the offsets might be good, but the surface may well oscillate in a non-fair fashion in-between the offsets.

In this particular problem, matching the

offsets of stations is not enough if the stations and the surface between the stations is not fair. Approaches that rely on interpolating splines suffer from this problem; it may be relatively simple to match the offsets on frames, but achieving the offsets together with a satisfactory overall shape can be considerably more difficult.

4. FITTING OFFSETS WITH NURBS CURVES

Among the different mathematical spline curves that exist, the presented method will use B-spline curves and surfaces, that can be easily exported to all the CAD programs, and are today a standard in the industry, not only the naval one.

To introduce the notation for the paper, a brief review of B-splines follows. A B-spline curve is formed by several pieces of polynomial curves, called Bézier pieces, and the whole curve is C^2 (common curvature or second derivatives) at the junctions in the case of cubic B-splines. The curve is defined with a polygon, called the control polygon, and with an interpolation algorithm that allows its construction to relate the curve to the control polygon. The interpolation steps are encoded in a family of piecewise polynomial functions, $B_j^n(u)$, called B-spline functions of n^{th} degree, and are calculated using De Boor's algorithm [5]. Cubic B-splines are the most widely used curves in ship design and the ones that generally fit better the traditional loftman's splines.

A B-spline curve, $s(u)$ in Eq. (1), is a linear combination of basis functions and $m + 1$ control points, P_j , as coefficients. Therefore, B-spline curves are parametric, $x = X(u)$, $y = Y(u)$, $z = Z(u)$, where the parameter u is considered inside $[0,1]$ in this paper. In the 2D plane, $P_j = (X_j, Y_j)$, $j = 0, \dots, m$, generate a B-spline $s(u)$ of the n^{th} degree,

$$\begin{aligned} s(u) &= \sum_{j=0}^m P_j \cdot B_j^n(u) = (X(u), Y(u)) = \\ &= \sum_{j=0}^m \left(X_j \cdot B_j^n(u), Y_j \cdot B_j^n(u) \right) \end{aligned} \quad (1)$$

Where the basis functions are obtained with the De Boor's algorithm of (2)

$$\begin{aligned} B_j^0(u) &= \begin{cases} 1 & u \in [u_{j-1}, u_j) \\ 0 & \text{otherwise} \end{cases} \\ B_j^n(u) &= \frac{u - u_{j-1}}{u_{j+n} - u_{j-1}} B_j^{n-1}(u) + \frac{u_{j+n+1} - u}{u_{j+n+1} - u_j} B_{j+1}^{n-1}(u) \end{aligned} \quad (2)$$

The basis functions, $B_j^n(u)$, depend on the knot vector, u_j . In this work, the knot vector has been selected to be uniform with a multiplicity equal to the order of the curve at its ends, where the order is defined as the degree + 1. In this manner, the B-spline interpolates the ends of its control polygon at $u = 0$ and $u = 1$, and it is tangent at its ends to the first and last segment of its control polygon. This last property simplifies the mathematical definition of the curves used in the method.

The mentioned reasons on section 3 do not suggest the use of an interpolating B-spline. Instead, an approximating curve is needed. For every station of index "d", the B-spline $c_d(u)$ will not pass through the data points exactly but will pass close enough to the points to capture the inherent shape. This is the well-known Least Squares (LS) approximation [6].

In this problem, $np+1$ data points Q_0, \dots, Q_{np} from the analytic curves will be approximated by a B-spline of p^{th} degree, with $N+1$ control points P_0, \dots, P_N , $N < np$ that are unknown and are obtained as the final result of the calculations. These points may adopt constrained angles at the ends of the stations, which is a normal practice in ship hull forms.

The general LS problem is described by the over-determined set of $np+1+2$ equations with $N+1$ unknown variables:

$$\begin{aligned}
B_0^p(t_0) \cdot P_0 + B_1^p(t_0) \cdot P_1 + \dots + B_N^p(t_0) \cdot P_N &= Q_0 \\
B_0^p(t_1) \cdot P_0 + B_1^p(t_1) \cdot P_1 + \dots + B_N^p(t_1) \cdot P_N &= Q_1 \\
\vdots &\vdots \\
B_0^p(t_{np}) \cdot P_0 + B_1^p(t_{np}) \cdot P_1 + \dots + B_N^p(t_{np}) \cdot P_N &= Q_{np} \\
B_0^p(t_0) \cdot P_0 + B_1^p(t_0) \cdot P_1 + \dots + B_N^p(t_0) \cdot P_N &= \text{tg}(\angle) \\
B_0^p(t_{np}) \cdot P_0 + B_1^p(t_{np}) \cdot P_1 + \dots + B_N^p(t_{np}) \cdot P_N &= \text{tg}(\angle)
\end{aligned} \quad (3)$$

Where B_i^p corresponds to the i^{th} basis function of a p^{th} degree B-spline that is calculated using de Boor's algorithm considering a uniform knot vector and t_j ($j=(0, n)$) represents the parameters associated to the data points. Matrix expressions are convenient to solve the problem:

$$\begin{bmatrix} B_0^p(t_0) & B_1^p(t_0) & \dots & B_N^p(t_0) \\ B_0^p(t_1) & B_1^p(t_1) & \dots & B_N^p(t_1) \\ \vdots & \vdots & \vdots & \vdots \\ B_0^p(t_{np}) & B_1^p(t_{np}) & \dots & B_N^p(t_{np}) \\ B_0^p(t_0) & B_1^p(t_0) & \dots & B_N^p(t_0) \\ B_0^p(t_{np}) & B_1^p(t_{np}) & \dots & B_N^p(t_{np}) \end{bmatrix} \cdot \begin{bmatrix} P_0 \\ P_1 \\ \vdots \\ P_N \end{bmatrix} = \begin{bmatrix} Q_0 \\ Q_1 \\ \vdots \\ Q_{np} \\ \text{tg}(\angle) \\ \text{tg}(\angle) \end{bmatrix} \quad (4)$$

$$[M] \cdot [P] = [Q] \Rightarrow [M]^T \cdot [M] \cdot [P] = [M]^T \cdot [Q]$$

This system of equations is solved by multiplying both sides of Eq. (4) by $[M]^T$, which creates a determined $(N+1)$ by $(N+1)$ linear system. This type of system can be poorly conditioned, especially if a large number of control points is used. A conventional technique should not be used to solve this ill-conditioned system. Instead, single-value decomposition of $[M]^T \cdot [M]$ and a later back-substitution process is performed. The solutions of this system are the control points of the best B-Spline fitting.

Approaching this problem with a standard parameterisation such as centripetal or chord-length is correct but does not consider the effect of the distance of the data points to the B-spline. In this method, we adopt a parameterisation based on a minimum distance. These parameterizations were first introduced in (5), but the way to obtain the minimum distance is different. The process is iterative

and is described by the following three steps:

1. The method starts with a centripetal parameterisation of the Q_i points and system (3) is solved. This produces a starting curve of the iterative process only for the first loop.

2. For each Q_i , the minimum distance to the B-Spline is calculated. This is done by dividing the B-Spline $c_d(u)$ in Bézier curves $b_j(u_L)$ ($j=1, N-p$) of the p^{th} degree and computing the minimum distance to the corresponding Bézier piece, leading to a solution to Eq. (5).

$$(Q_i - b_j(u_L)) \cdot (b'_j(u_L)) = 0 \quad (5)$$

This equation is solved in the local domain of the Bézier curves. Because the equation is a polynomial equation, $u_L \in [0,1]$ and specific algorithms for this type of equation can be used. These algorithms do not require an initial guess, which would be required if a Newtonian method were used in the B-Spline domain. The current method uses a Jenkins-Traub 3-stage algorithm [7]. The valid solution will be a non-complex solution of $u_L \in [0,1]$.

Once the solution has been found, the local u_L for the Bézier domain is easily converted into its global value t_i in the B-spline domain. This t_i value is the parameter associated with the point Q_i when solving system (3).

3. After obtaining the t_i ($i=1, np$) values, the distance $d_i = (Q_i - c_d(t_i))$ is computed, which is the Euclidean distance between Q_i and the B-Spline. This distance is used to check the shape requirement. If the maximum distance d_i ($i=1, n$) is above a given tolerance, steps 2 and 3 are repeated until an acceptable maximum distance is achieved. More specifically, the quality of the obtained curve is measured using the tolerance constraint and the shape of the B-Spline is amended using parameterisation (5). An example of the effect of the iterative parameterization can be seen on Fig. 3

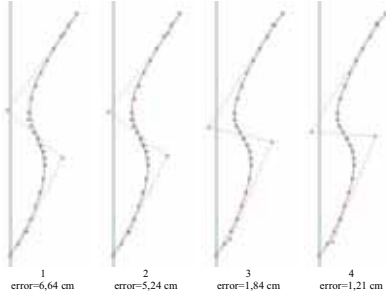


Fig. 3 Effect of the iterative parameterization

If the tolerance is not obtained in fewer than 50 iterations, then the number of control points $N+1$ has to be increased. The increment of the degree p in this procedure can also reduce the maximum distance. However, increasing the number of control points has a more substantial effect and a higher degree raises the complexity and the computational time.

When the fitting is finished for all the stations, a set of q B-splines with the form of Eq. (7) are obtained. Fig. 4 presents an example of the described fitting (because of the symmetry of ship stations, curves displayed at the left are stations between the aft end and the middle length of the ship, while curves at the right are taken from the middle towards the fore end).

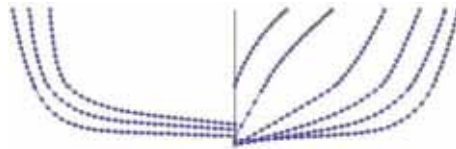


Fig. 4 Example of the constrained fitting

This particular example shows a cubic B-spline fitting where all the curves have 9 control points and have been obtained after 30 iterations. The maximum distance from the data points to the curves is less than 2 mm, and the dimensions of this ship are 6690 mm x 2000 mm. The tolerance can be improved by increasing the number of control points as needed to satisfy construction requirements.

5. LOFTING SURFACE OF THE STATIONS

The last step of the method is to create a B-spline surface that leans on the stations previously defined. The generalisation from curves to surfaces is not difficult thanks to the properties of B-splines and a lofting surface of the stations can easily be defined. The transition from spline curves to spline surfaces is achieved by turning the control polygon into a control net of control points $W_{ij}(X_{ij}, Y_{ij}, Z_{ij})$ using the same B-spline basis for the two parameters u and v , as well as using two different lists of knots $\{u_1, \dots, u_{N+n}\}$, $\{v_1, \dots, v_{M+m}\}$.

The lofting process of a set of q B-splines (stations) with the same degree and list of knots is as follows: find a B-spline surface S with degree n by m and $(N+1)$ by $(M+1)$ control points and a list of knots $\{u_1, \dots, u_{N+n}\}$ and $\{v_1, \dots, v_{M+m}\}$ according to Eq. (6) that interpolates q different B-splines c_d ($d=0, \dots, q-1$) of n^{th} degree with $N+1$ control points and a list of knots $\{u_1, \dots, u_{N+n}\}$ with the form of Eq. (7).

$$S(u, v) = \sum_{i=0}^N \sum_{j=0}^M \mathbf{W}_{ij} \mathbb{B}_i^n(u) \mathbb{B}_j^m(v) \quad (6)$$

$$\mathbf{c}_d(u) = \sum_{i=0}^N \mathbf{V}_{id} \mathbb{B}_i^n(u) \quad (d=0, \dots, q-1) \quad (7)$$

Note that \mathbf{V}_{id} are the control points of the different stations obtained from system (4) and expressed in a matrix form. Values for N , M and q depend on user preferences regarding the definition of the surface. The interpolation can be written as:

$$\begin{aligned} S(u, v_d) &= \sum_{i=0}^N \left(\sum_{j=0}^M \mathbf{W}_{ij} \mathbb{B}_j^m(v_d) \right) \mathbb{B}_i^n(u) = \\ &= \sum_{i=0}^N \mathbf{V}_{id} \mathbb{B}_i^n(u) = \mathbf{c}_d(u) \quad (d=0, \dots, q-1) \end{aligned} \quad (8)$$

This group of equations has to be solved for a set of values of parameter v_d ($d=0, \dots, q-1$),

termed the choice of the parameterisation. The centripetal parameterisation produces good results for ship hulls, and the spacing between stations should not be very different to avoid wiggles. By identifying equal coefficients for every row of Eq.(8), $i = 0, \dots, N$, the following linear system is obtained:

$$\sum_{j=0}^M \mathbf{W}_{ij} [\mathbf{B}_j^m(v_d)] = \mathbf{V}_{id} \quad (d=0, \dots, q-1)$$

To obtain a unique solution for this system, $M+1 = q$, where q is the number of stations that defines the ship hull. The $(M+1)$ by $(N+1)$ solutions are the control points \mathbf{W}_{ij} of the lofting surface of Eq. (6) containing the stations.

6. EXAMPLE

The method is applied to a sailing ship of 12.5 m length of Fig. 5. Once that the offsets of the stations (left part of Fig. 5) have been obtained, the B-spline fitting of section 4 is applied, considering cubic B-splines with $N=8$ control points. The curves obtained are shown in Fig. 6, where the control points corresponding to the bow stations are showed. Tolerances of the fitting are below 0.01 m.

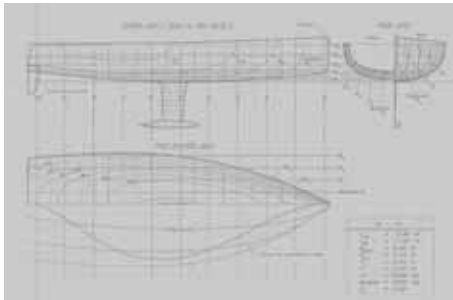


Fig. 5 Lines drawing of a sailing ship

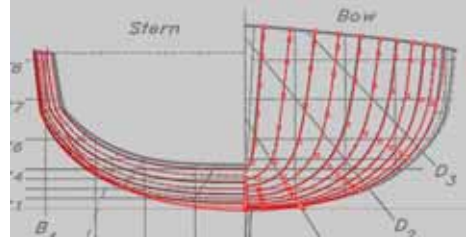


Fig. 6 B-spline fitting of the stations

Finally, the lofting process explained in section 5 is applied and a final B-spline surface is obtained of degrees 3×3 is applied. The control points and a rendered view of the surface are showed in Fig. 6.

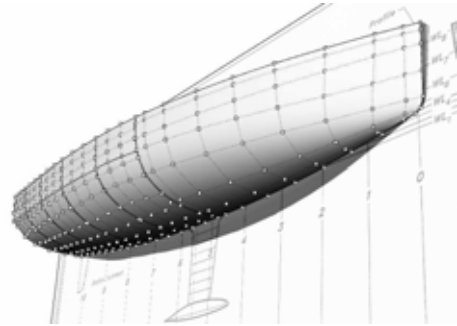


Fig. 7 Control points of the lofting surface

7. CONCLUSIONS

This paper has presented a method to model ship hulls from its 2D lines drawing, into a 3D NURBS surfaces definition.

Offset points from the stations are merged with a B-spline using least square fitting, considering a distance tolerance into its parameterisation. This has an important influence from the construction point of view. The number of longitudinal and transverse control points on the surface can be defined, so the definition of the surface can be increased and then manually modified in later stages to include appendages or other design requirements, which may be difficult to include in the initial definition of a constraint design

procedure.

The final result is a B-spline surface that can be easily exported to specialised computer programs for further stages of the design process. The same hull representation can be used for the design and in all subsequent stages of the process of developing the ship.

modeling, parametric ship design and geometric modeling.

REFERENCES

- [1] L. Cao, J. and Liu. 3D reconstruction from line drawings without hidden lines. IEEE Trans FAMI, 30 (3):507-517, 2006
- [2] Y. Leclerc and M. Fischler. An optimization-based approach to the interpretation of single line drawings as 3D wire frames. IJCV, 9(2):113-136, 1992.
- [3] L. Cao, J. and Liu. 3D Object retrieval using 2D line drawing and graph based relevance feedback. Proc. ACM International Conference on Multimedia, 105-106, 2006
- [4] X. Chen and S. Kang. Sketching reality: realistic interpretation of architectural designs. ACM Trans. Graph., 27(2):1-15, 2008.
- [5] G. Farin. Curves and Surfaces for CAGD, Ed. Morgan Kaufmann, San Francisco. 2001.
- [6] L.A. Piegl and W. Tiller. The NURBS Book, Ed. Springer, p. 410-413, 1997.
- [7] A. Jenkins and J. Traub. A three-stage variable-shift iteration for polynomial zeros and its relation to generalized Rayleigh iteration, Numerische Mathematik 14(3): 252-263, 1970.

ABOUT THE AUTHORS

1. Dr Francisco L. Pérez Arribas is Professor in the Naval Architecture and Marine Engineering School of Madrid (ETSIN), UPM. He received BS and MS degrees in Naval Architecture from the Universidad Politécnica de Madrid. He has working experience at the ETSIN Towing Tank in CAD and CASHD. His current research interests include ship hull

MODERN BOOK ON GEOMETRY

Vladimir SHELOMOVSKII, Svetlana NOSULYA
Russia

ABSTRACT: The authors have been solving geometric problems and teaching mathematics for more than twenty years. The gained experience have been used to create the concept of modern book on geometry which uses visualization, maximum interactivity, gives the opportunity to work with 3D-images. The corresponding system based on DGS GInMA and providing e-book or textbook on geometry have been created. It was implemented in the form of the textbook "Advanced Geometry" for Russian students, partially translated into English. This interactive textbook is destined for students who don't like formulae, but experience great interest in images and their dynamic transformation. The paper presents a small part of GInMA electronic textbook with visual library and interactive solutions, the topic "Sejfriedian". Dynamic geometric illustrations make it convenient to introduce the Sejfriedian study in training courses on geometry. The interactive textbook is available on [7]. Download freely distributed DGS GInMA [8] and e-book, install the software and get to enjoy the interactive drawings by clicking on the Figures in the text.

Keywords: Interactive textbook, visual library, 3D geometry, DGS GInMA, step-by-step solution.

1. INTRODUCTION

The authors have been solving geometric problems and teaching mathematics for more than twenty years. The gained experience have been used to create the concept of modern book on geometry which uses visualization, maximum interactivity, gives the opportunity to work with 3D-images. The corresponding system based on DGS GInMA and providing e-book on geometry have been created. It was implemented in the form of the textbook "Advanced Geometry" for Russian students, partially translated into English. This interactive tutorial with visual library and interactive solutions have been developed for students who don't like formulae, but experience great interest in images and their dynamic transformation.

The paper presents a small part of GInMA electronic textbook, topic "Sejfriedian", which has been chosen for several reasons. This topic is very extensive, many-sided, it organically combines seemingly different themes: the singular points of a triangle, simedians, polar

lines, projective transformations of the line onto itself, the plane onto itself, and much more. This material is the result of the original scientific research, which have been obtained by the authors with the use of DGS GInMA and formed the content of the corresponding tutorial. Dynamic geometric illustrations make it convenient to introduce the Sejfriedian study in training courses on geometry. We already have the experience of "Sejfriedian" tutorial use in education in Advanced course of geometry for gifted students.

The interactive textbook is available on [7]. Download freely distributed DGS GInMA [8] and the e-book, install the software and get to enjoy the interactive drawings by clicking on the Figures in the text.

2. GINMA E-BOOK REPRESENTING A RESULT OF SCIENTIFIC RESEARCH

Scientific research in geometry frequently demand observation of the complicated geometric configuration in dynamics. This allows to no-

tice some new regularities, which may be overlooked when we use classic computational methods. Such new regularities become the statements and even the new theorems after the proof. Dynamic visualization also helps to search the proof.

Ensuring the accuracy of calculations and constructions, not conceding the accuracy of the numerical programs, DGS GInMA provides visualization of the geometric constructions with the opportunity to work with 3D-images. This is particularly useful in the calculation of integrals and in the calculation of the volumes of solids of intersection. Difficulties arise when the program builds the plot and has to choose some of the solutions. Especially often such problems arise at the intersection of quadrics.



Figure 1: Sphere crossed by the ellipsoid.

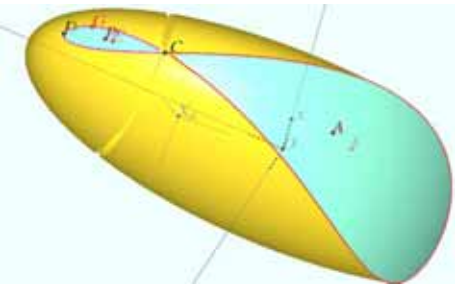


Figure 2: Solid of intersection.

When the solutions in some cases are real in other cases they are imaginary. An example of such a situation is shown in Figures 1 and 2. The solid of intersection of the blue sphere and the yellow ellipsoid is considered in the case when the line of intersection has a self-intersection point.

GInMA supports convenient means of transformations (inversion, symmetry, dilation, rotation), convenient means of surfaces construc-

tion, interactive moving of the points on the surface, numerical construction of a surface. As an example, Figure 3 shows the study of the Meissner solid which is important to explain the idea and technique of the lemon-shaped surface.

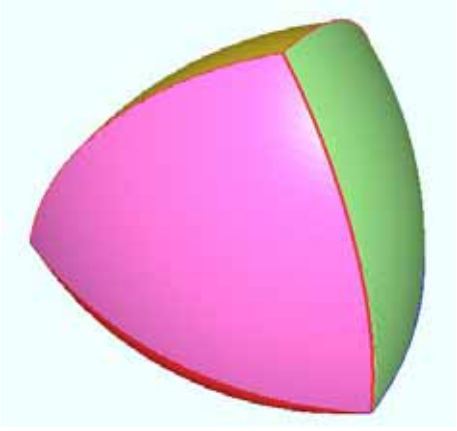


Figure 3: Meissner solid with the lemon-shaped surface.

DGS GInMA is effectively used by geometers thanks to its convenience for research, providing wide opportunities for hypotheses testing, for the objects properties checking (the existence of common points of contact and intersection and so on).

In their mathematical research the authors apply DGS GInMA at all stages of the study. We perform geometric construction, view it and observe some regularity. It is important to make sure that this regularity persists when the parameters change. DGS GInMA meets this need and allows us to carry out further investigation of geometrical configuration in dynamics. The study typically needs to quickly perform different transformations. We perform the transformation, look at the resulting construction, check it numerically. The program displays the numerical calculations on the screen.

Figure 4 shows an example of inversion. The sphere is tangent to three given spheres. Under the inversion two given spheres are mapped into the parallel planes.

to the original text of the problem being solved. The corresponding system for GInMA e-book have been created. Interactive visual library provides the transition from the text or the picture to the corresponding interactive file. The initial HTML-file of GInMA textbook contains the links to the Manual and video assistants and offers a large selection of geometric topics.

Click on the name of the Topic takes us to the PDF file with the topics text with its own topics Library. The topics Library is formed as a series of small icons, convenient to cover the subject at a glance. Each icon shows the characteristic Figure of a problem. For example, Figure 6 shows the library of the topic: "Volumes of solids of revolution" which allows us to get acquainted with different methods, in particular with the Golden-Popp method, to explore the solid of revolution obtained by rotation of a triangle or a circle, to explore equivalent replacements of the figure of rotation or the rotation axis.

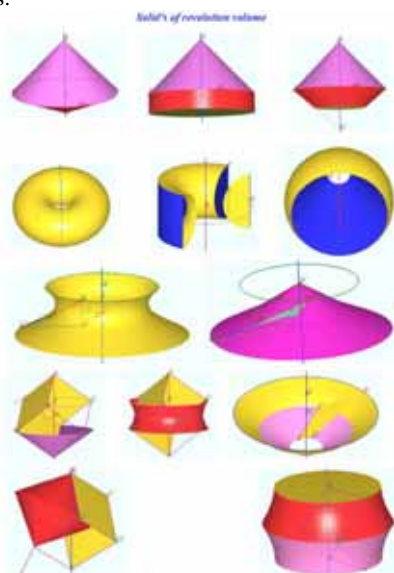


Figure 7: Library of the topic: "Volumes of solids of revolution"

Click on the Library icon takes us to the problem. Here we see the text of the condition,

the text of the solution with the Figures, recommendations on the research, hyperlinks to the theorems necessary for the solution. Click on the Figure in the text takes us into the interactive file, where the static Figures turns into the dynamic drawings.

The interactive files provide detailed step by step solutions. Here we may change the image, pass by the steps of the interactive solution, perform the additional construction. For example, in each interactive file of the topic: "Volumes of solids of revolution" we may interactively change the figure of rotation and investigate the configuration in dynamics. Videos of some solutions you may see on YouTube, channel member Vladimir Shelomovskii. Just easily, we may come back from the interactive file into the problem text.

GInMA e-book "Advanced Geometry" has a large number of interactive files containing information on a wide range of topics, such as polar lines, special points of a triangle, projective transformations, inversion, and much more. During the educational process students are reviewing many textbook files, some of the files are quickly viewed, some files are carefully studied. Students pick up material for themselves, which is necessary for their currently learning.

The author practices the following methodology of GInMA textbook use in teaching. At the lesson the lecturer announces to the students that the next lesson will be devoted to a certain topic. Students are preparing at home using GInMA textbook, by selecting and studying the theme files which they need. In the next lesson they get the tasks on this topic. The problem or the question may be shaped differently than the problems from textbook, but the essence is the same. Students when solving at the lesson may use GInMA textbook. Having prepared in advance, with understanding of the material, they can find a similar tasks, useful information among the GInMA tutorial files.

The tasks are considered the most difficult on mathematical competitions which require to perform additional construction or to go beyond

the conditions in the construction. The difficulty is caused, in addition to the complexity of the task, also by the properties of the human psyche. It is often psychologically difficult to go beyond the usual. Students get used to do it in dynamic research with the use of GInMA textbook in every lesson. Thus they overcome the psychological barrier which hinders "to go beyond".

Dynamic visualization of spatial geometric constructions is sometimes necessary not only for students but even for teachers. Difficulties arise when solving problems, for example, with skew lines. In the projection the skew lines may look like intersecting lines. Some teachers told the author, how they have tried to find the point of intersection over and over again when solving problems in the case of skew lines, where it did not exist. When using GInMA textbook in everyday teaching practice the teachers don't confront with such difficulties.

The author's experience as a teacher shows that interactivity and visualization help to gain a deeper understanding of geometry. Students easier memorize theorems presented in a dynamic form, in the form of visual image, in which students themselves change, move the elements. GInMA textbook provides this capability.

4. GINMA E-TBOOK SAMPLE: THE TOPIC "SEJFRIEDIAN"

As an example, in the paper we consider a part of the electronic GInMA e-book, the topic "Sejfriedian". It is the result of scientific research, which have became the content of the corresponding tutorial. It is worth noting that already at the stage of research DGS GInMA became an important instrument of scientific work. Further the result of the research have became the content of GInMA textbook, the topic "Sejfriedian". The topic "Sejfriedian" is interesting due to the fact that it is very extensive, many-sided, combines seemingly different themes: the singular points of a triangle, simedians, polar lines, projective transformations of the line onto itself, the plane

onto itself, and much more. Dynamic GInMA illustrations make it convenient to introduce the Sejfriedian study in training courses on geometry for the development of students geometric vision.

Base configuration. We use the Sejfriedian in the form shown in Figure 8. Suppose that the line l with points A, B and C be given. We may draw a circle and place on it the points K, L, M, K_1, L_1 , and M_1 . We assume that we may build the straight line EQD so that all the points from each group from the groups of four points $(L, L_1, A, E), (L, K_1, Q, C), (L, M_1, D, B), (M_1, K, Q, A), (K_1, M, D, A), (L_1, M, Q, B), (K_1, K, E, B)$ lie on the corresponding straight lines. The object of exploration is the construction of such configuration, its properties investigation and its applying for Steiner mapping in some special cases.

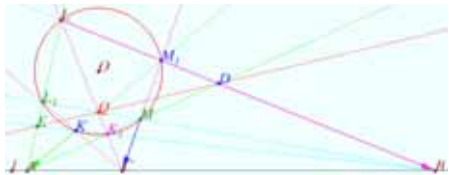


Figure 8. Sejfried configuration.

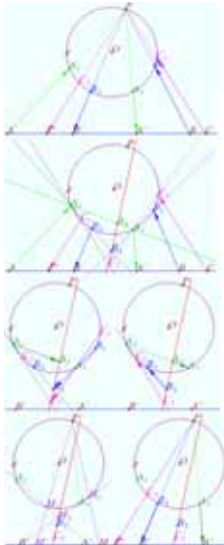


Figure 9. Steiner solution.

Steiner solution. Let the straight line l and the points A, A', B, B', C , and C' be given. We want to create the projective transformation mapping the straight line into itself such that the point A is mapped into A', B into B' and C into C' . This problem was solved by Steiner [1-4]. The solution is shown in Figure 9.

Problem for the author and further for students The solution accomplishes the desired transformation in cases where a couple of lines A_1B_3 and A_3B_1 (A_1C_3 and A_3C_1) intersect at a single point. Therefore, the Steiner solution does not exist in special cases, for example, when we map $A \rightarrow B, B \rightarrow C, C \rightarrow A$ or $A \rightarrow A, B \rightarrow C, C \rightarrow B$. The problem is convenient to be discussed in the education process because it can be solved in different ways and detects deep geometric relationships. One method of solving is the use of projective transformations. Figure 10 shows some non trivial configurations.



Figure 10. Projective transformations.

Another approach is based on the properties of singular points of the triangle (the Lemoine points, the Fermat points, the Apollonius points), polar lines and symmedians, and on the applying the inversion transformation. Fragments of corresponding libraries are shown in Figures 11 and 12.

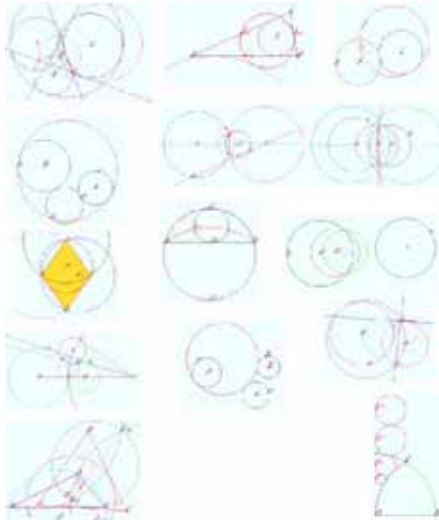


Figure 11: Inversion.

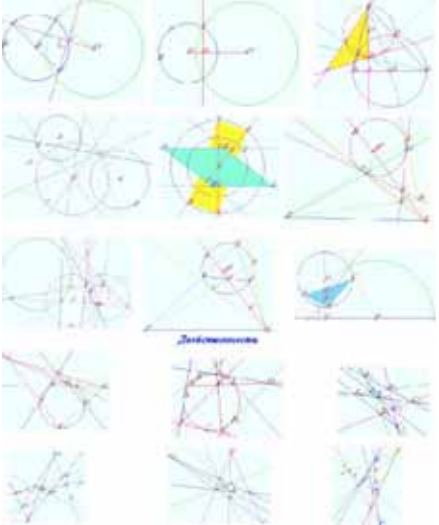


Figure 12. Polar lines.

5. CONCLUSIONS

Modern book on geometry which provides visualization, maximum interactivity, opportunity to work with 3D-images have been created based on DGS GInMA. Such e-book may be useful for representing the scientific research and for geometry teaching.

REFERENCES

- [1] V.Shelomovskii. *Theorems and constructions created using DGS GInMA.*, at the Proceedings of the 15th International Conference on Geometry and Graphics (ICGG 2012), pp 688-697, 2012. ISBN 978-0-7717-0717-9.
- [2] M. Seifried, V.Shelomovskii. *Elementary Proof of Seifriedian Properties.*, at the Proceedings of the 17th Asian Technology Conference in Mathematics (ATCM 2012), pp 342-352, 2012. ISBN 978-0-9821164-4-9.
- [3] 1. Jacob Steiner. *Die Theorie der Kegelschnitte, gestützt auf projectivische Eigenschaften.* Bearb. von Dr. Heinrich Schröter. 2 Aufl, Leipzig 1867.
- [4] Jacob Steiner. *Gesammelte Werke*, Bd, I,II, 1882. (Геометрические построения, выполняемые с помощью прямой линии и неподвижного круга. Перевод Д.М.Синцов, 1939).
- [5] *Steiner's Theorems on the Complete Quadrilateral.* Jean-Pierre Ehrmann, Forum Geometricorum, Volume 4 (2004) 35–52.
- [6] Yiu P. *Introduction to the Geometry of the Triangle.* Chapter 10. General conics. [<http://www.math.fau.edu/yiu/geome-try.html>]
- [7] <http://deoma-cmd.ru/files/documents/Seifriedian.pdf>

SOFTWARE PACKAGES

- [8] GinMA Software, <http://deoma-cmd.ru/en/Products/Geometry/GInMA.aspx>

ABOUT THE AUTHOR

1. Vladimir Shelomovskii is Associate Professor, Ph.d in mathematics, General Director and leading specialist in developing of math e-books, Deoma LLC. Twenty years he has devoted to high-tech technology, led the research group of the Russian Academy of Sciences engaged in the development of aerodynamic windows for gas lasers. He has been working in education for twenty five years as the math lecturer, the Chair of department of mathematical modeling and math methods in economics at the Murmansk University. His research interests include solving of difficult mathematical problems, logic, DGS development and geometry teaching with its use.
2. Svetlana Nosylia is member of the DGS development team. Her research interests include solving of logic problems, computer design and geometry teaching with the use of DGS.

THE MONGE THREE POINT SPACE RESECTION PROBLEM

Riccardo MIGLIARI¹, Federico FALLAVOLLITA² and Marta SALVATORE³

^{1,3}Sapienza University of Rome, Italy ²Alma Mater Studiorum University of Bologna, Italy

ABSTRACT: The study illustrates the solution of the three point space resection problem, treated by Gaspard Monge in Section V of *Leçons de Geometrie Descriptive*.

The problem entails the construction of the intersection curves of three tori. To solve this problem, Monge introduces several simplifications but, nevertheless, makes a mistake; this mistake has already been pointed out by Gino Loria regarding the number of solutions allowed by the problem [11]. The mathematical representation, thanks to its high level of accuracy, today permits not only an efficacious solution of the general case, it also highlights without difficulty the right number of solutions.

We applied this theory to a case of photogrammetric rendering, difficult to carry out by means of the tools offered by commercial software. Case in question concerns the reconstruction of the architectural volumes, now lost, which were located along the road that crosses a village, near Rome. As is known, the reconstruction of points in space from two images is possible if these images are projective and we have at least two projective orientated stars. The first image is a vintage photograph (1892), the second image is a surveyed plan of the masonry still present at the site. Therefore, one of the two projective stars is assimilated to a class of vertical straight lines. With regard to photography, the problem is articulated in two typical phases of photogrammetric processes: internal orientation and absolute orientation. For the absolute orientation we used the pyramid vertex method, in use since the Eighteenth Century, which consists in determining the projection center from three given points of which the positions in space are known.

The solution to the problem posed by the case study can be considered as a useful result. More interesting, however, is the result of the intersection of the three tori with the incident axes (fig. 1). It is, in fact, a graphic process that Gaspard Monge had already proposed in 1798 as a suitable alternative to a system of equations that he considered difficult to solve. In particular, Monge explains how the descriptive geometrical procedure, involving the vision of the represented forms, allows you to exclude in a simple and direct manner the solutions that resolve the problem from theoretical point of view, but do not solve it in the real case because they lead to unrealistic placements of the projection center. Thus, the symbiosis between calculation and analog description, Monge had predicted in these words: «[...] la géométrie descriptive porteroit dans les opérations analytiques le plus compliquées l'évidence qui est son caractère, et, à son tour, l'analyse porteroit dans la géométrie la généralité qui lui est propre [...]» [18].

Keywords: Descriptive geometry, three point resection problem, Gaspard Monge, photogrammetry.

1. INTRODUCTION

Much has been said and written about the role of Gaspard Monge in the history of representation, sometimes attributing him the 'invention' of descriptive geometry, sometimes re-

ducing, perhaps unfairly, his merit to that of editor of a book already written in the course of previous centuries. But, beyond this difficult evaluation, there is no question that Monge must be recognized as given scientific dignity to the representation of objects in three dimen-

sions. This possibility, distinctly expressed from Frézier as a requirement of engineering and architecture, was not, before Monge, expressed equally explicitly, even in the work of François Sylvestre Lacroix, that also preceded, of course, the publication of the Monge lectures. In his effort to theorize and give general validity to the proceedings hitherto confined to stereotomy and gnomonics (i.e., in special cases), Monge sought a synthesis between the analytical study of geometric shapes and their properties and the graphic and visual representation, therefore, of the same forms and properties, in the belief that the combination of the two methods could give new impetus to the geometry, « [...] Il seroit à désirer que ces deux sciences fussent cultivées ensemble [...] » [18].

Curiously, two centuries after the golden age of *Géométrie Descriptive*, Monge's hope seems to be realized in the mathematical representation of the digital universe. In fact, as we have noted several times, the accuracy that the calculation gives digital representation, the possibility of building forms in three dimensions, the possibility of using curved and double-curved surfaces as tools in geometrical construction, combined with the power of visual analogy, allow us to experiment the Monge's far-sighted vision [13].

This paper traces, using the synthetic method, Monge's solution, to the problem of the pyramid vertex construction, given its triangular base data and the corners at the top, and showing an original application to the reconstruction of volume reconstruction of a partially lost agglomerate of historical interest.

2. MONGE AND THE THREE POINT SPACE RESECTION PROBLEM

The IV section of the *Géométrie Descriptive* of Gaspard Monge, entitled *Application de la méthode de construire les intersection des surfaces courbes à la solution de diverses questions*, is devoted to the solution of six questions, which concern the determination of the position of a point starting from that of other geometric entities [16].

Three of these questions are of theoretical nature and three instead derive from practical problems. In particular, the fifth and sixth questions concerning the construction of the position of a point on a slope, notes that it is the relative positions of the other three points, which are visible from the point sought [11]. In the first case, the tools available are a goniometer and a plumb line, in the second case it is the goniometer only; therefore, we consider the second case, which generalizes the problem.

The point sought, where we imagine the observer was placed, corresponds to the pyramid apex that has its basis in a triangle whose vertices are the three known points on the ground. From his position, the observer can measure, the angles that the projecting lines passing from the three given points, form between them. The problem then consists in determining the position of the pyramid vertex given its base triangle ABC and the angles at the vertex α , β e γ .

We construct the three corners of the triangle AB , BC , AC ; let us consider one of this, for example, AB , and then we consider the α vertex angle subtended by AB .

We construct a circle on the ABC plane capable of the angle α said. This circle is divided into two arcs from the AB chord. The $\pi - \alpha$ angle of the AFB arc is supplementary to the α angle of the ADB arc. Then, rotate the two arcs of circle around the AB axis and thus generate the double-nappe surface of a torus; we call inner nappe, one generated by the AFB arc, and outer nappe, the one generated by the ADB arc. All points belonging to the two nappes of this torus are seen by the AB segment, under the same angle, α for the outer nappe and $\pi - \alpha$ for the inner nappe (fig. 1).

We reiterate the procedure for the AC and BC edges of the pyramid base, and we get the other two tori of double nappes, the locus of points seen by the data segments in the respective β and γ corners.

The pyramid apex, by construction, must belong to the surfaces of the three tori, in particular their outer nappes, with respect to the

example proposed in the figure. These nappes intersect each other in three skew curves (one or two branches). The intersections of these

curves give rise to a maximum of eight points, eight vertices of a pyramid which can satisfy the conditions imposed by the problem. These

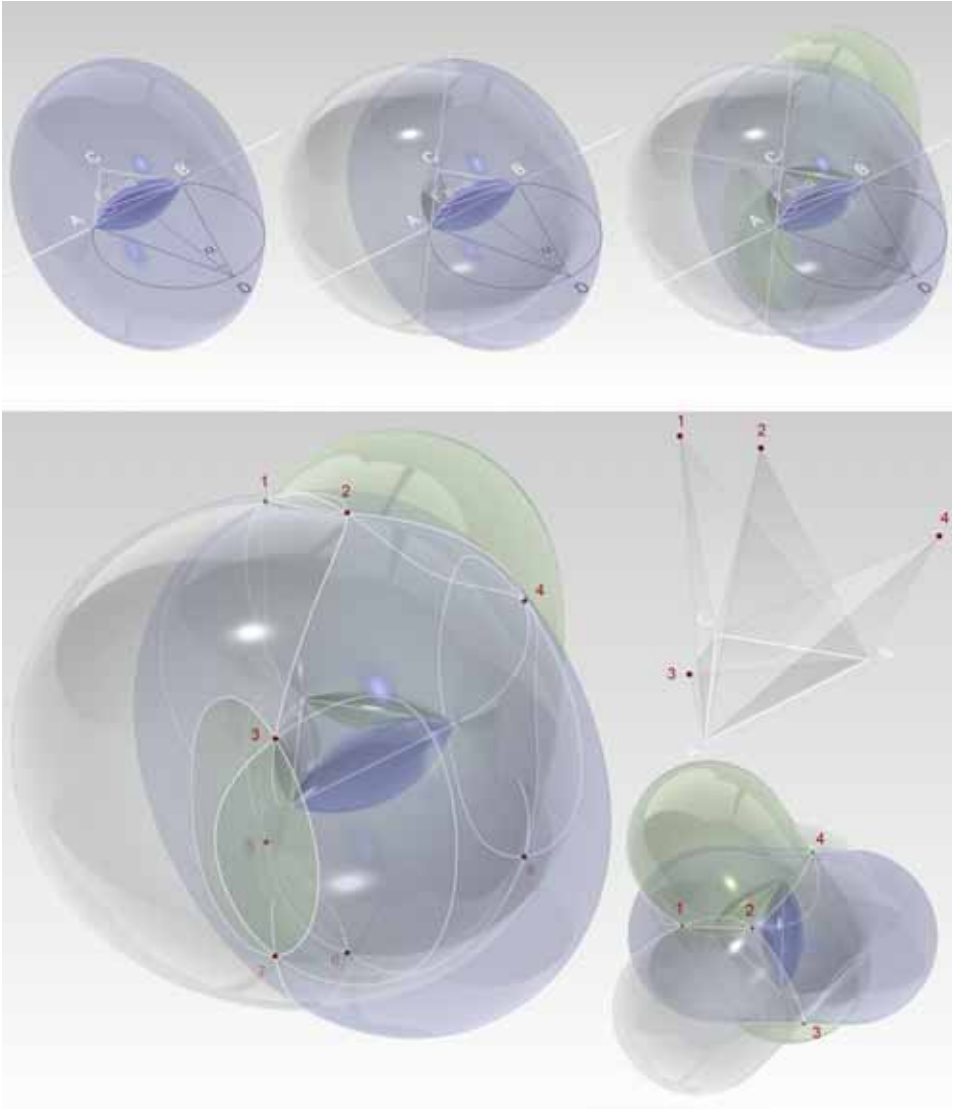


Figure 1: The Monge three point space resection problem, digital synthetic solution.

solutions are all valid from a theoretical point of view, but only one satisfies the problem of a practical nature that we set. Since the eight solutions are symmetrical with respect to the **ABC** plane, we can exclude the half, in particular those which are found beyond the **ABC** plane of symmetry. Among the remaining four, one meets the conditions of the observer's position relative to the plane **ABC**. The synthetic method allows you to directly select the right one.

This question, which has had a significant impact in solving the problems of orientation in photogrammetric surveys, stimulated, from its first publication, the interest of several mathematicians. Among these, Lacroix and Hachette, who dedicated various articles to this problem, published in his major works of descriptive geometry [8-9-10], but also Nichola Fergola and students of his school [10], including in particular Vincenzo Flauti [6-7]. There were many ambiguities on the number of possible and admissible intersections for the solution of this problem, presented for the first time by Lagrange in the *Acts of Berlin* of 1773. Monge himself, while recognizing eight feasible solutions, sixty-four inferred from the calculation, arrived at a general solution that first Flauti, then Loria reported as incorrect [11-6]. Hachette a few years later re-examined the issue and identified 16 feasible solutions [8], then reduced to 12, in a subsequent article [10] (fig. 2). Hachette, however, never specifies the maximum number of admissible solutions. In fact, the Hachette's solutions refer to the configurations of the problem he represented, and include vertices which contextually observe the given angle and the supplementary angle and therefore, must be discarded. Vincenzo Flauti, in *Geometria di sito sul piano e nello spazio*, generalized the question by explaining that twenty solutions are found, derive from the combinations, three by three, of the six nappes of the three tori [6]. Of these twenty, twelve do not solve the problem, because they are derived from the intersection of the outer nappes with inner nappes of the same torus and therefore

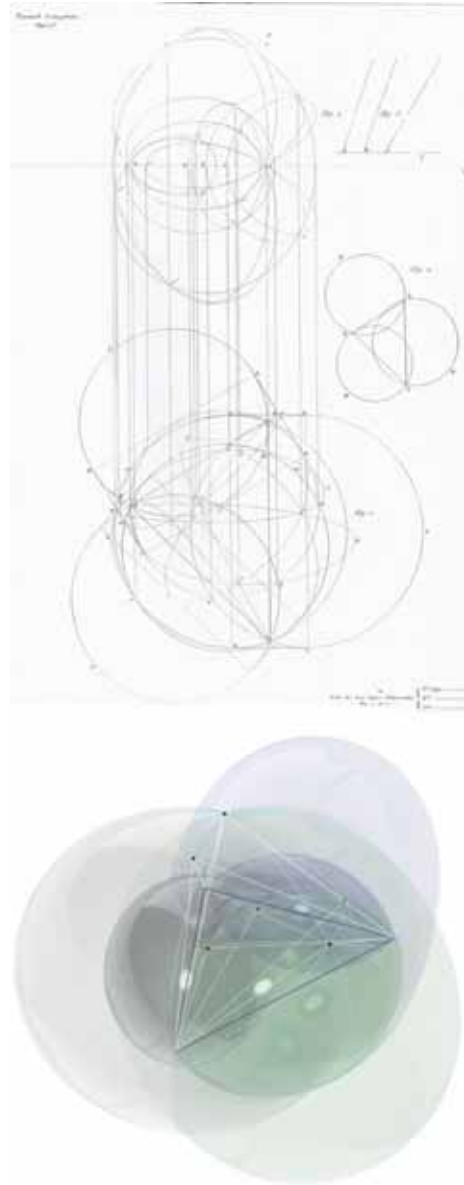


Figure 2: Hachette's solution, from the *Traité de géométrie descriptive*, 1822, planche C.

must be discarded. Therefore eight intersections resolve the problem.

In all the solutions described it is not adequately clear which method identifies the intersections that solve the problem in its generality. It should be noted that this problem is placed at the limit of the possibilities of the synthetic method, when using graphical 2d methods, such as representations of orthogonal projection performed with great skill first by Monge, then by Hachette. Today, the digital synthetic method allows us to immediately track the number of intersections, while in the past, the same 2d graphical method may have given rise to some ambiguities, such as those reported in the Hachette's text referring to his figures.

3. AN INTERESTING CASE

The procedure described above by Monge finds an interesting application in a particular case illustrated below. This is a reconstruction of the volumes of a paper mill in the village of Grottaferrata, Italy, today lost as a result of an earthquake, from some survey data: a vintage photo from 1892, a picture of AOS 1938 and the relief of the walls still present on the site.

In general, the reconstruction of points in space, from their two-dimensional images, is possible provided that the images are of projective nature and provided that you have at least two projective-oriented stars. In the present situation, the first image is made from a vintage photo (fig. 3), while the second can be reconstructed from an aerial photo of 1938, and by the relief of the remaining walls. Regarding the survey, you can assimilate the star to a class of vertical straight lines, and the same can be done, with good approximation, to the aerial photo, considering: the portion of the flight of AM on the one hand, and the fact that we are considering only a small portion of the frame, on the other. Thanks to the survey, this star may also be considered oriented. Regarding the first image, however, i.e. the photograph of 1892, the problem is more complex and must be broken down into two phases typical of photogrammetric



Figure 3: Vintage village photo and 3:2 format.

procedures: the internal orientation (position of the center of projection with respect to the frame) and the absolute orientation (frame position in space at the time of shooting). The solution to these two problems permits us the full reconstruction of the projective star and therefore the reconstruction of the location of points in space that no longer exist in reality.

In the photogrammetric restitution, as in any other activity of a scientific nature, you can follow proceedings of deterministic character or probabilistic character. In the first case, having available reliable data and sufficient conditions we arrive at the result without uncertainty. In the second case, when the data are not entirely reliable and conditions are insufficient, you can still achieve a convincing result, through an iterative process, in which the hypothesis that integrates the given data, are introduced throughout the process and the results they provide are compared with the results of previous iterations. In this process we discard the hypothesis that generate, in the iteration, the worst results and keep those that provide the best results, thus converging towards the solution. In this case it is not possible to apply a deterministic process, because the data are insufficient. It is possible, however, to reconstruct the projective star and its orientation in space by successive approximations, converging towards a result corroborated by the reduction of the error. It is, possible, to formulate a

hypothesis, and comparing it with the results process. The presence of the building volumes still intact in fact, permit us to orient the star and check that the largest possible number of points present in the photograph coincide with the corresponding real points. The hypotheses that will be chosen with respect to others, are those which, in testing, give the best results and the fewest errors. Remember that the interior orientation is known when you know:

- the principal point, i.e. the point at which the camera optical axis intersects the plane of the photographic plate;
- the focal length, i.e. the distance of the projection center (entrance pupil of lens) from the plate.

If the photograph was printed in its entirety, that is, without the cuts that have altered the

shape of the plate, and if you have not applied optics and mutual translations or rotations of the plate (as can happen in some professional cameras) the principal point is in the intersection of the diagonals of the frame. In this case, the original format was altered in the press, as can be seen considering the upper left corner of the image, the edge of which is broken. It can therefore be assumed that the section of the higher edge belongs to the original plate, while the lower portion is the result of a cut that was done to approximately align the horizon and the vertical to the edges of the print.

Regarding the relationship between the sides of the original plate it is not easy to formulate a hypothesis because the cameras of the time, and the craftsmanship, used various formats. The archives of Count Primoli (1851 - 1927), for example, are made up mostly of square plates, while many cameras of the time used

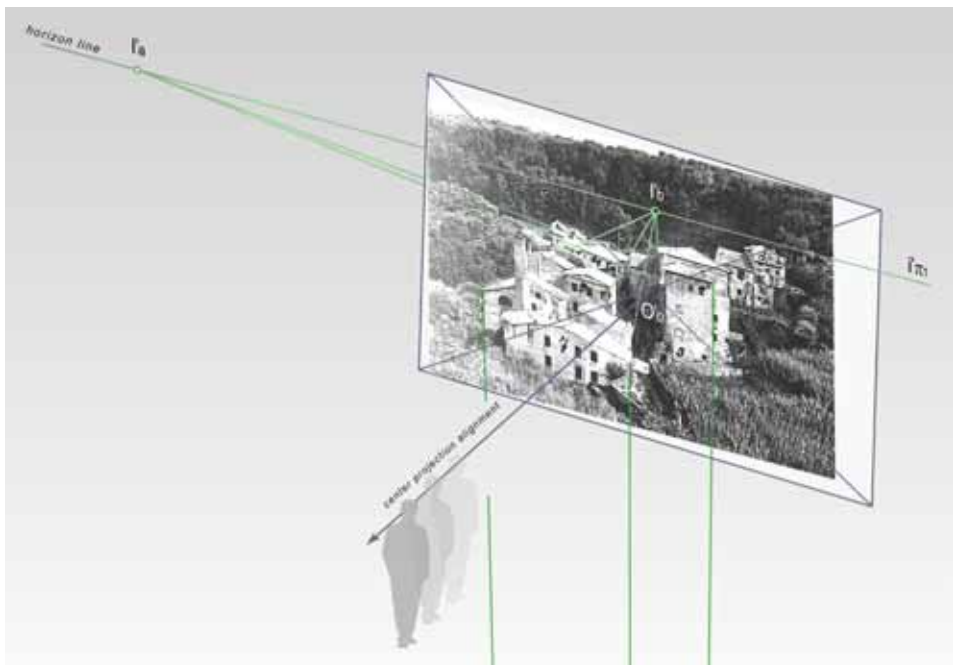


Figure 4: Reconstruction of the horizon line and the principal distance direction.

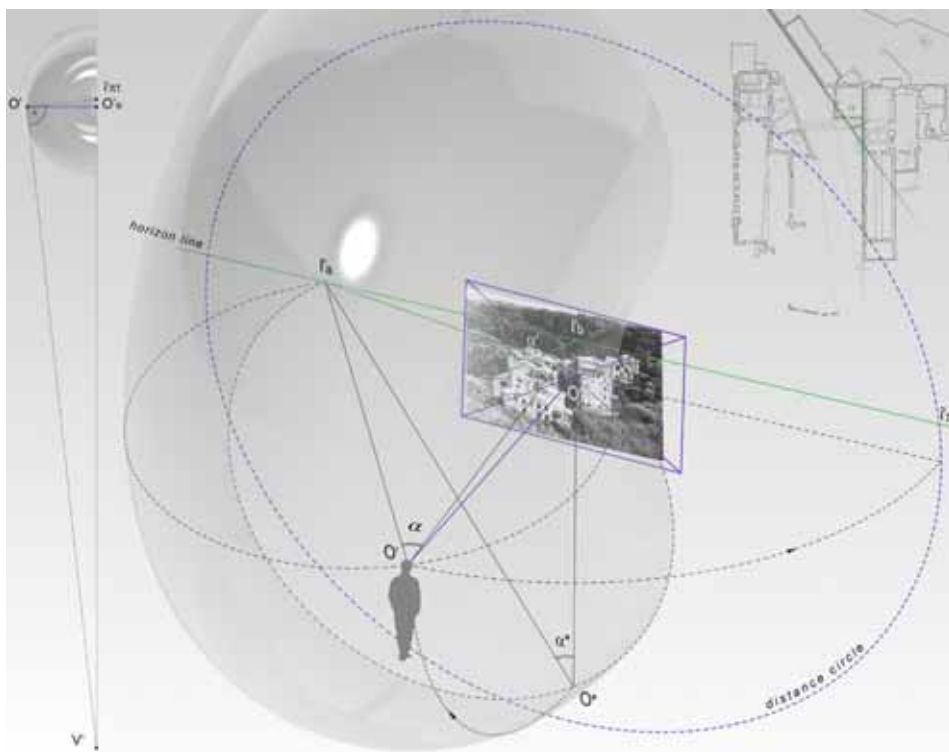


Figure 5: The frame internal orientation.

plates whose sides are in the ratio 3:2. The picture we are looking at has sides that are in the ratio 23.9:16.8, which is different from both the 4:3 ratio (i.e. 24:18), both from the 3:2 ratio (i.e. 24:16). Therefore, both hypotheses were verified. We illustrate here the one that has led to better results.

A rectangle was built whose sides are in the ratio 3:2. The left side coincides with the short segment of the original plate which is seen in the upper left. The other sides of the rectangle are such as to include the entire image. We can therefore formulate the hypothesis that the principal point is in the intersection of the diagonals of the rectangle constructed as above. If so, the center of projection will be on the

perpendicular to the plate plane and will pass through the principal point (fig. 3).

The cameras of the time used for focus, a sled that allowed a wide range of the lens in front of the plate. Therefore, the focal length cannot be established a priori, even in an approximate way. You can, however, use the knowledge of some corners. In fact, the projection center must be such as to see the direction of the straight line and the arrangement of the buildings plans, from the angle they form in reality.

The most important of these angles is what the plumb direction line makes with the horizontal position. In photography, the direction of the vertical projects into the vanishing point of

the vertical edges of buildings and can be found with good approximation in point V' . Reconstruction of the horizon is not so sure, because the plate is angled, and therefore the horizon does not pass through the principal point, and because it is also hard to find pairs of straight horizontal lines and vertical lines. You can, however, be sure that you will find the horizon above the buildings, and below the limit which separates the woods behind the sky (fig. 4). The checks carried out on various assumptions included in this part of the perspective plane lead to the adopted solution.

You may also consider the angle formed by the two buildings of greatest importance: which include the church (A) and what comprises the turret (B). This angle, deduced from the survey, is equal to thirty-four sexadecimal degrees (or its supplement). Therefore, we construct on the horizon the images of the directions of horizontal lines that belong to the two planes. Then we can construct the locus of points in space capable of these angles, which is a surface of revolution (torus). The projection center is the point where this surface meets the principal distance (fig. 5). This process is repeated until you find the position of the center that satisfies both conditions. The focal length is determined in relation to the size of the frame. In the hypothesis that the original plate measure 6x9, the focal length is equal to 151 mm, length that is compatible with the optical construction of the time, being equivalent to that of a 60 mm lens on the Leica format.

The pyramid vertex method is used for the absolute orientation. The position of the projection center coincides with the apex of a pyramid that sees three points still present on the site, location unknown, under known angles. The instrument used for the angular measurement is, in this case, a vintage photograph that, once oriented, is capable of restoring the angles formed by the straight lines that coincide with the points photographed. The control points, instead, were chosen on the basis of survey data, rather distant from each other, to minimize error. In particular, E' , I' and L' were chosen,

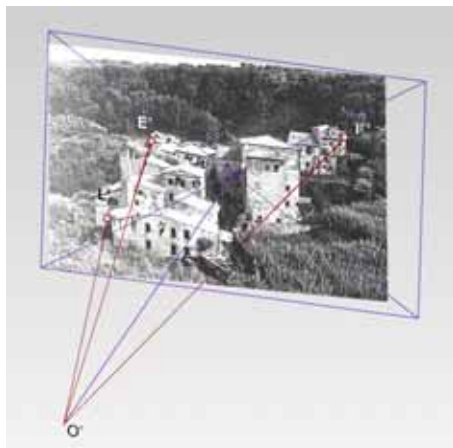


Figure 6: The three points of the pyramid base.

seen from the projection center under the following angles expressed in sexadecimal degrees (fig. 6):

$$E'OI' = 17.7155$$

$$I'OL' = 23.7762$$

$$E'OL' = 8.7568$$

The solutions of the problem are found through the intersection of geometrical loci of the points that see the data three point below the measured angles (fig. 7). In this case there are, in all, four options, three of which have a purely theoretical significance because they are the vertices placed below ground and the opposite vertex of observer site. The center of the projective star generated by the photographic shooting lies instead in the vertex that represents a possible and coherent position with the image (fig. 8). The procedure for absolute orientation is completed by placing the system formed by the plate and the projection center in the center as constructed above and imposing the collinearity of the edges of the pyramid with the three straight lines connecting the projection center O to the points E' , I' and L' , which are the photographic image of the real ones E , I and L (fig. 9).

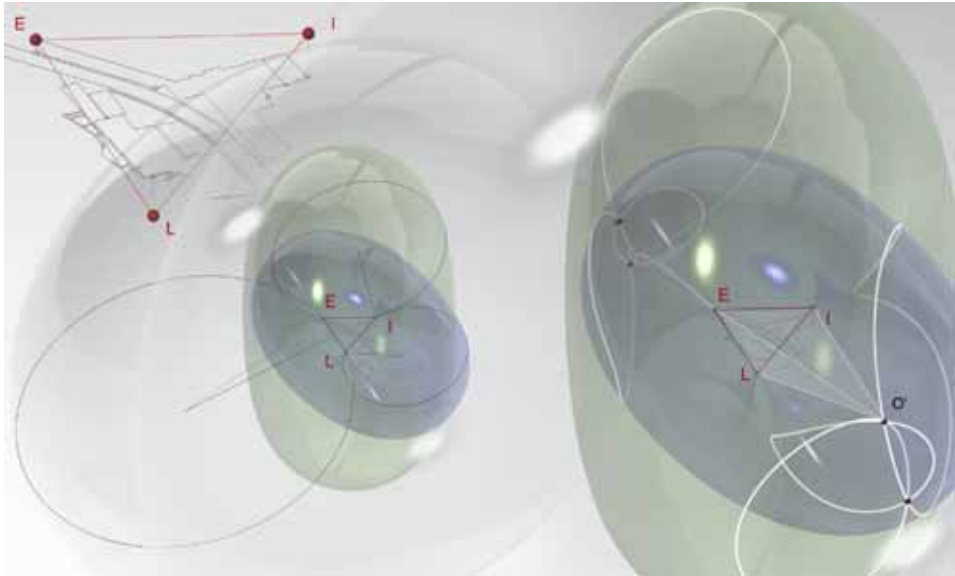


Figure 7: The three point space resection problem applied to the case study.

Note that performing this operation, the system also requires a precise orientation in space and therefore the straight line projecting from the center in the vertical direction must be vertical, or at least very close to verticality. In the present case the straight line form with the vertical an angle of 1.6 degrees sexadecimal. The residual errors can be calculated collimating these points. In this case we obtained errors of decimeter magnitude.

4. CONCLUSIONS

The solution to the problem posed by the case study can be considered a useful result. More interesting, however, is the result of the intersection of the three tori with the incident axes. It is, in fact, a graphic process that Gaspard Monge had already proposed in 1798 as a suitable alternative to a system of equations that he considered difficult to solve. In particular, Monge explains how the descriptive geometrical procedure, involving the vision of the represented forms, allows one to exclude, in a

simple and direct manner, the solutions that might resolve the problem from a theoretical point of view, but do not actually resolve it, because they lead to unrealistic placements of the projection center. Thus, the symbiosis between calculation and analog description, Monge had predicted in these words: «[...] *la géométrie descriptive porteroit dans les opérations analytiques le plus compliquées l'évidence qui est son caractère, et, à son tour, l'analyse porteroit dans la géométrie la généralité qui lui est propre* [...]» [18].

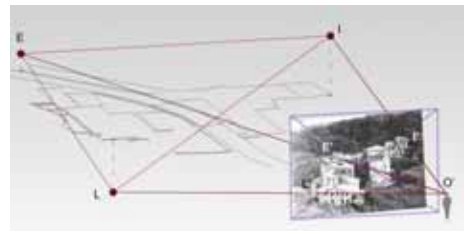


Figure 8: The frame absolute orientation.

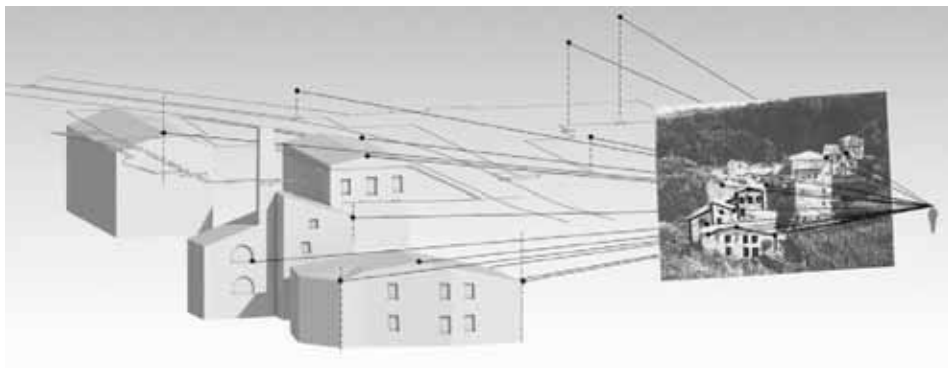


Figure 9: The village volumes reconstruction.

REFERENCES

- [1] V. Cardone. *Gaspard Monge: scienziato della rivoluzione*, CUEN, 1996.
- [2] M. Carpicci. Un nuovo modello grafico-analitico per l'orientamento di fotogrammi mediante l'intersezione di tre tori, in *Emergenza Rilievo 2°* vol. Kappa: Roma, 2001. pp.224- 252.
- [3] P. Dore. *Fondamenti di fotogrammetria - Fototopografia da terra e da aerei*. Zanichelli: Bologna, 1938. I, 4, p. 33.
- [4] C. Dupin. *Essai Historique Sur Les Services Et Les Travaux Scientifiques de Gaspard Monge*, Bachelier: Paris, 1819.
- [5] F. Fallavollita, M. Salvatore. A digital synthetic method, in *Geometria's 14*, proceedings of XIII Encontro da Aproged, Lisboa, 17-18 may 2014, pp. 181-190.
- [6] V. Flauti. *Geometria di sito sul piano e nello spazio*, Stamperia privata dell'autore, Napoli, 1842, pp. 178-189.
- [7] V. Flauti. *Raccolta di opuscoli matematici*, Stamperia privata dell'autore, Napoli, 1851, pp. 178-189.
- [8] J.N.P. Hachette. *Corr spondence sur l' cole Royale Polytechnique*, vol II, Klostermann: Paris, 1813, pp. 332-337.
- [9] J.N.P. Hachette. *Suppl ment de la G om trie descriptive de Monge*, Klostermann: Paris, 1812, pp. 110-118.
- [10] J.N.P. Hachette. *Traite de G om trie descriptive*, Corby: Paris, 1822, pp. 151-160.
- [11] G. Loria. *Storia della geometria descrittiva*, Hoepli: Milano, 1921, pp. 114-115.
- [12] G. Loria. *Nicola Fergola e la scuola de matematici che lo ebbe a duce*, Genova, 1892, pp. 66-72.
- [13] R. Migliari. *Geometria descrittiva*, Novara: Citt Studi, 2009.
- [14] R. Migliari, F. Fallavollita. The geometric fundamentals of design: towards a new descriptive geometry. In P. Zsombor- Murray, A. Sprecher, B. Angeles (edited by), *The 15th International Conference on Geometry and Graphics*, Montr al, Canada, August 1-5, 2012, pp. 473-482.
- [15] R. Migliari. The Apollonian problem and Descriptive Geometry. In *Disegnare idee immagini - Ideas Images*. Roma: Gangemi, 2008, vol. 36, pp.22-33.
- [16] R. Migliari. Descriptive Geometry: From its Past to its Future. In *Nexus Network Journal*, vol. 14, n.3, Springer, Birkh user Verlag Basel, 2012 pp. 555-571.
- [17] G. Monge, *G om trie descriptive*, Boudoin: Paris, 1798 (ed facsimile Jacques Gabay, Sceaux 1989), pp. 16, 95-102.

- [19] M. Piazzolla - Beloch. *Elementi di Fotogrammetria Terrestre ed Aerea*. Cedam: Padova, 1934. Parte II, V, 34, p 55.
- [20] R.M. Haralick, C.N. Lee, K. Ottemberg, M. Nölle, (1994). Review and Analysis of Solutions of the Three Point Perspective Pose Estimation Problem. In *International Journal of Computer Vision*, 13, 3, Kluwer Academic Publishers: Netherlands. 1994. pp. 331-356.
- [21] R. Taton. *L'oeuvre scientifique de Monge*, Presses Universitaires de France: Paris, 1951.

ABOUT THE AUTHORS

1. Riccardo Migliari is Full Professor in Fundamentals and Applications of Descriptive Geometry since 1990 at the Faculty of Architecture at the 'La Sapienza' University in Rome. He is assiduously engaged in research, particularly in the areas of Descriptive Geometry and of Representation and Instrumental Survey of Architecture. He directed, as Scientific Manager, the architectural survey of Coliseum in Rome during the preliminary studies for the restoration of the monument undertaken by the Archaeological Superintendence of Rome. From year 2003 he deals, in particular, with the renewal of the studies on the scientific representation of the space, within in the evolutionary picture of the descriptive geometry, from

the projective theory to the digital theory and from the graphical applications to the digital modeling. <http://www.migliari.it>

2. Federico Fallavollita, Architect, is Assistant Professor at the Department of Architecture at the University of Bologna. PhD in Science of Representation and Surveying. He is engaged in research, particularly in the areas of Geometry and Representation and Instrumental Survey of Architecture. He is especially interested in problems of projective and solid geometry through the digital synthetic method, problems associated with the use of digital tools in architectural drawing and architectural survey:

<http://www.unibo.it/faculty/federico.fallavollita>

3. Marta Salvatore, Architect, is PhD in Survey and Representation of Architecture and the Environment. She is engaged in research in the areas of Descriptive Geometry, Representation and Instrumental Survey of Architecture. In particular, she directs her research activities to the geometry, its history and renewal through the synthetic method, by digital tools of representation. She is author of several articles on the subjects, available on her website:

<http://www.martasalvatore.it>

MULTIVARIANT DATA IN COMPLEX QUIZ LAYOUT FOR COMPUTER ENGINEERING GRAPHICS

Daiva MAKUTĖNIENĖ

Vilnius Gediminas technical University, Lithuania

ABSTRACT: The problem in this paper is inspired by information visualization technologies and difficulties in assessment of students' knowledge. Information visualization has been defined as "the use of computer-supported, interactive, visual representations of abstract data to amplify cognition. The task is to create a complex quiz layout by applying information visualization methods and by analyzing numerical and graphical datum in application for graphics examination.

Visualization can be a powerful tool for analyzing data and presenting results in education. Application may be based on relational database system and CAD system for execution analysis.

Relational database systems are flexible to organize, store, and manipulate any numerical and textual data, CAD system has an ability of testing tasks solution.

Testing on geometry cannot replace all variety of control forms of students' knowledge. Nevertheless, test tasks allow not only leading the general current control to reveal mastering computer graphics knowledge, but also to execute training and correcting functions. Visualization is a powerful tool for analyzing data and presenting results not only in science, engineering and medicine. This paper reviews ways in which it can be used in distributed and/or collaborative environments in education.

Keywords: information visualization, graphical tasks solutions, multiple-users, collaborative environment in education.

1. INTRODUCTION

The problem in this paper is inspired by information visualization technologies and difficulties in assessment of students' knowledge. Information visualization has been defined as "the use of computer-supported, interactive, visual representations of abstract data to amplify cognition" [2]. The task is to create a complex quiz layout by applying information visualization methods and by analyzing numerical and graphical datum in application for graphics examination. This paper deals with techniques in which data visualization can be used in distributed and/or collaborative environments. Visualization can be a powerful tool for analyzing data and presenting results in education. Application may be based on rela-

tional database system and CAD system for execution analysis.

2. CONTROL AUTOMATION IN GRAPHICAL EDUCATION

The focus on control automation of graphic knowledge in the teaching process was even before the origin of computers. In Lithuania "programmed" survey was used for this matter in a form of question templates (Figure 1) and evaluated with the help of response templates. Separate course (usually one lecture by volume) had several prepared options of test cards, where students had to answer to the questions either "yes" or "no". Although it is limited, but the structuring of the material proved to be useful later.

<div> <div>3-28</div> <div> ВИСТ ВІСЛІДОВІ КАТЕДРА КАФЕДРА ЧЕРЧЕННЯ ВІСІ </div> </div>		АЛГАНТИ-СЕРВІС		
КЛАСИФИКАТОР		1	2	3
1. Які геометричні фігури, які геометричні тіла (циліндр, конус, куля) можна побудувати за допомогою черчівних інструментів?		$\overline{A_1}$	$\overline{A_2}$	$\overline{A_3}$
2. Які геометричні фігури, які геометричні тіла (циліндр, конус, куля) можна побудувати за допомогою черчівних інструментів?				
3. Які геометричні фігури, які геометричні тіла (циліндр, конус, куля) можна побудувати за допомогою черчівних інструментів?		$\overline{A_1}$	$\overline{A_2}$	$\overline{A_3}$
4. Які геометричні фігури, які геометричні тіла (циліндр, конус, куля) можна побудувати за допомогою черчівних інструментів?		<i>Horizontalizacija</i>	<i>Frontalizacija</i>	<i>Profilizacija</i>
5. Які геометричні фігури, які геометричні тіла (циліндр, конус, куля) можна побудувати за допомогою черчівних інструментів?		<i>В перспективі</i>	<i>В ортографі</i>	<i>В перспективі</i>
6. Які геометричні фігури, які геометричні тіла (циліндр, конус, куля) можна побудувати за допомогою черчівних інструментів?				

Figure 1: Old fashion cards of the programmed survey

Technical development possibilities let a programmed survey equipment appear (Figure 2), where students could mark answers to the test questions not on a template sheet, but pressing the appropriate button on the panel placed in the table. The results could be seen on the teacher's table.



Figure 2: Old-fashion programmed survey equipment

Of course, these technical tools let some stages of knowledge control to automate only partially. Only beginning of personal comput-

ers and comfortable programming languages allowed a new type of control options to develop, which enable service not only fragments of knowledge, but also means of graphic work control. Computer techniques in graphics are used for decades, but level of control automation is low enough. The control process requires attention, the results may have important consequences. It is still associated with technical problems (screen size, flicker) in computing. Control may not only be unfair, but also troublesome, so automation is desirable.

Despite improvement of control methods, they can not be applied in computing. Only the control principle remains. People were always seeking for rational work, assessing technical feasibility of the times. Computing gives unlimited opportunities, only people lack desire to implement. Practically all human work can be done by computer. Often much faster and more qualitatively. Engineering graphics problems scale (the same as engineering activities themselves) is very wide (geometric drawing, projection drawing, graphics application, etc.) therefore authentication opportunities of geometric drawing only are overviewed.

Integrating machine learning and information visualization is potentially fruitful. [1]. Process of perception is more efficient when human information analyst (lecturer) working with data and graphics engine in optimal way. This is accomplished in Figure 3.

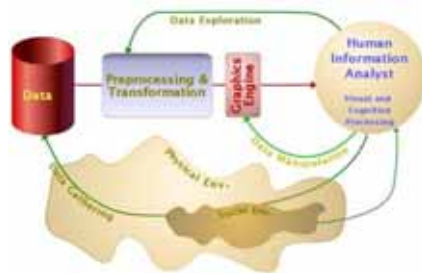


Figure 3: The diagram of process of visualization [2].

Geometric drawing consists of drawing lines and text. Various schemes of graphic work control are possible. VGTU teachers/professors

selected the one when automatically-formed task template is given for control and graphical data is compared with the result of the work [5,6]. In this case, two phases are distinguished in information processing model (Figure 4): task preparation and management and process control organization.

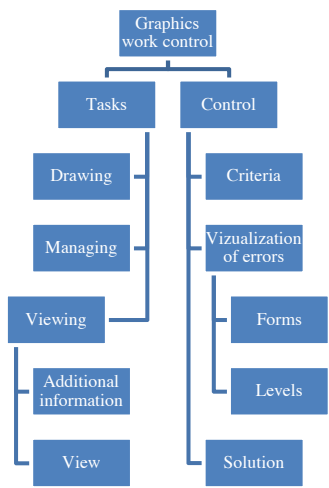


Figure 4: Information processing model

Task unit consists of task preparation functions, image making management and presentation to a student. Presentation must be provided with textual information to explain workflow-related knowledge. Control solves the questions of the analysis and evaluation.

Task-making problems and practice are analyzed and solved in academic work by Department of Engineering Graphics. Graphic version of the task is presented in Figure 5.

Task preparation is a huge work which requires much receptivity. Program tasks acquired certain advantages over paper-based tasks because a certain group of students perform work of the same complexity but different graphical appearance (Figure 5).

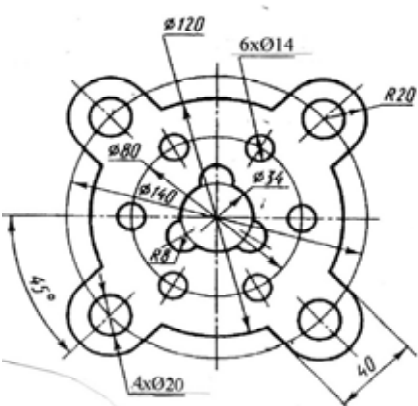


Figure 5: The old fashion task for examination.

Parametric tasks give us opportunity to apply computing possibilities more completely by automating work control. Graphic elements database of the drawing template is being formed during task presentation (visualization) [7].

After the drawing is done, graphic data base of the work is being formed. Control base is being formed by using elimination method (drawing elements are deleted from the list of template elements).

Control base elements are evaluated by using the criterion scale and conclusion is formed. In the initial stage it is appropriate to take advantage of the interactive screening criteria by adapting them in your computer version. The errors in the drawing are shown in graphically color scale (red - rough, yellow – slight, green - proofing), keeping on the criteria of the scale (Table 1).

Table 1. Graphical evaluation of errors

Error	Grade off	Comment of evaluation
Accuracy of the drawing	10	Inaccurate computer drawing is impossible, not allowed
Layers		
Layers are used incorrectly or missing at all	0,1x ER (ER - amount of er-	Commented errors

	rors)	
Axes		
Not enough or too many axes	0,1x ER	indicated errors
Axes of invalid length	0,1	Axis should be extended for 3-5 mm beyond the designated element.
Width of lines	0,1	Comment
Dimensions		
No dimensions	3	Comment
Dim style is prepared incorrectly	0,1...0,5	
Dim style is created in a correct manner but not applied	0,2	
Lack/too many dimensions	0,1x ER	indicated errors
R or \square are marked incorrectly	0,1x ER	R is marked in the place of \square and v.v.
Lack of note when repetitive elements appear or note is done incorrectly	0,1x ER	5x15 \square e.g.
axial and dimensional lines duplicate	0,1	Comment, which usually happens by adding dimensions which designate the distances between centers of circles
Dimensions are arranged incorrectly	0,1	Too far, too close, cross other elements when it can be avoided, etc.
Time		
Main time 60 min	0	
70 min	1	The optimum time exceeded 10 minutes
80 min	2	The optimum time exceeded 20 minutes

Usually one teacher and some students are involved in the training process. In order learning process would be successful, both sides need to know about the quality of the work carried out. Visual control alone is not enough because we can get more information from computer drawing than drawing on a sheet of paper. To automate computing drawing control, at least the following key parameters are

used:

- work accuracy (joints, grating, typical pixel mapping);
- Isolation of control parameters (repetitive elements, spam, etc.)
- Isolation of control area (window, layer, drawing, line type);
- Methods of control visualization (color, flicker, conclusion);
- Links of work adjustment and methods to provide conclusions.

Work accuracy is a multiple (accuracy of seamless connectivity, accuracy of elements deletion, accuracy of typical pixel mapping). Accuracy of the work is controlled by splitting a screen into windows where you can see automatically magnified parts of the controlled drawing (Figure 6).

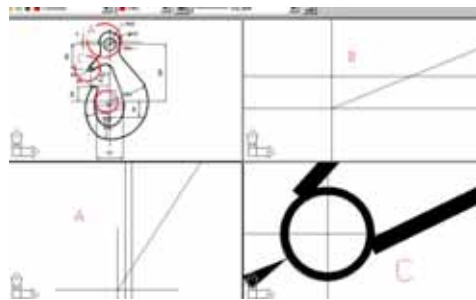


Figure 6. Control windows of accuracy in AutoCAD:

a, b - repetitive elements, c – results of inaccurate deletion

The errors are found and classified, the list of them and evaluation of the work are given to the user (Figure 7).

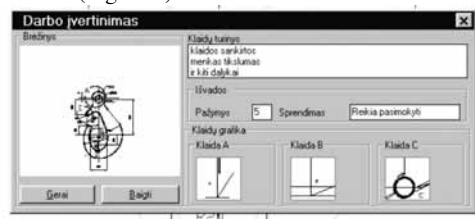


Figure 7. Form of presenting the results Automated control can play a dual role: to

facilitate and make more objective teacher's work and enable students to self-test their abilities. Evaluating computer work in a traditional way not only creates additional obstacles, but also minimizes computing capabilities. Reference drawings, which automatically become a computer image, have all premises to evaluate automatically (only process is needed to be modeled) and develop software.

Final mark of exam F in Engineering graphics department is sum of marks of homework (HW) and control work (CW), as in the following Equation (1):

$$F = 0.3(0.3HW + 0.7CW) + 0.7E \quad (1)$$

Relational database systems are flexible to organize, store, and manipulate any numerical and textual data, CAD system has an ability of testing tasks solution. The problem is a collaboration which allows multiple users at team environment to take part in the visualization process at different levels from viewing of images to the shared control of visualization methods, as in Figure 8.

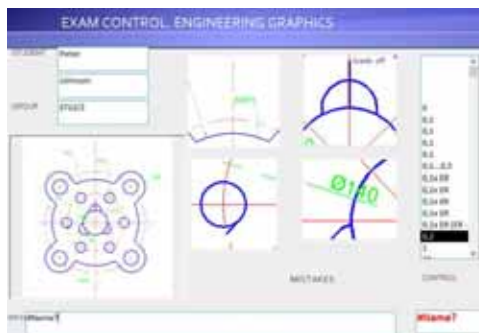


Figure 8: Possible exam control window for teacher.

Testing on geometry cannot replace all variety of control forms of students' knowledge. Nevertheless, test tasks allow not only leading the general current control to reveal mastering computer graphics knowledge, but also to execute training and correcting functions.

3. CONCLUSIONS

Visualization is a powerful tool for analyzing data and presenting results not only in science, engineering and medicine.

It is some ways in which it can be used in distributed and/or collaborative environments in education.

Maybe this can be solution how to improve students works testing.

Those problems can be good material for master studies works.

ACKNOWLEDGMENTS

The author gratefully acknowledge the support provided by Lionginas Čiupaila, assoc. prof. of Department of Engineering graphics and his master students. I am grateful for the anonymous reviewers on an earlier version of this article.

REFERENCES

- [1] C. Chen, Information Visualization: Beyond the Horizon, 2nd ed., Springer, 2004.
- [2] C. Ware. Information Visualization: Perception for design. 2nd ed., Morgan Kaufmann, 2004.
- [3] C. Romero, S. Ventura. Educational data mining: A survey from 1995 to 2005. *Expert Systems with Applications*, 33: 135–146, 2007.
- [4] Stuart T. Kard, Jock D. Mackinlay, Ben Scheiderman. Readings in Information Visualization, Using vision to think. Morgan Kaufmann Publishers Inc. San Francisco, CA, USA, 1999.
- [5] Astrauskas, Egidijus; Čiupaila, Lionginas. Geometrinės braižybos užduočių parengimo modeliavimas. Inžinerinė ir kompiuterinė grafika: konferencijos pranešimų medžiaga. Kaunas: Technologija, 2010.
- [6] Čiupaila, Lionginas; Zemkauskas, Jonas. Efficiency of problematic modelling and adaptation in engineering disciplines applying information technologies methods. 6th AECEF Symposium on Civil Engi-

neering Education in Changing Europe: proceedings. Vilnius: Technika, 2008.

- [7] Makutėnienė, Daiva; Čiupaila, Lionginas; Zemkauskas, Jonas. Taikomosios inžinerinės grafikos kurso modeliavimo ypatumai. Inžinerinė ir kompiuterinė grafika: konferencijos pranešimų medžiaga, Kaunas: 2012.

ABOUT THE AUTHOR

Dr. Daiva Makuteniene is associated professor in Vilnius Gediminas technical university, Department of Engineering graphics at Faculty of Fundamental Sciences since 1991. E-mail: delta@vgtu.lt.

NAVIGATING MULTI-DIMENSIONAL LANDSCAPES IN FOGGY WEATHER AS AN ANALOGY FOR GENERIC PROBLEM SOLVING

David RUTTEN

Robert McNeel & Associates, Austria

ABSTRACT: Developing algorithms that solve specific problems is a major area of research in computer science. Minimizing runtime complexities, reducing memory overhead, quickly returning intermediate or partial results, and balancing speed vs. accuracy are all important issues that need to be well understood when one wants to master algorithm design. It is safe to say that the people who need algorithms vastly outnumber those who can write them. However if we are willing to sacrifice performance, the application of *generic solvers* may democratize the field.

In principle, generic solvers are capable of solving most problems. This isn't quite as magical as it sounds, mostly because it involves a large amount of preparatory work by intelligent human beings. This includes defining the problem *phase space* and the relevant *fitness function*, which together make up the *fitness landscape*. Generic solvers are designed to find their way around these landscapes and to converge on high ground as quickly as possible.

This paper serves as an introduction to the theoretical side of generic problem solving with a strong focus on the geometry and topology of fitness landscapes. Along the way, several implementations of popular solver algorithms will be discussed.

Keywords: Generic Solvers, Phase Space, Fitness Function, Fitness Landscape, Peak Finding, Hill Climbing, Simulated Annealing, Evolutionary Algorithms

1. PROBLEM SOLVING

The problem of how to systematically solve problems goes back a long way. The history of *algorithms* can be traced back to the mathematicians *Brahmagupta* and *Abū 'Abdallāh Muḥammad ibn Mūsā al-Khwārizmī* from the 7th century India and 9th century Persia respectively. Algorithms are sets of unambiguously repeatable instructions which, when confronted with the same question, reliably give the same answer. The earliest known algorithms were devised to solve specific arithmetic problems such as multiplication or finding greatest common divisors.

Only in the last few centuries have mathematicians begun to approach algorithms from an analytic point of view, creating an *arithmetic of algorithms* which allows for their systematic classification and evaluation.

1.1 Problem categories

Computational mathematicians have come up with a set of classes to categorize different types of problems. Some of the better known classes are *P*, *NP*, *NP-Complete*, and *NP-Hard*. The class to which a problem belongs indicates roughly how quickly it can be solved. For example, it is always possible to generate answers to *P* class problems within a reasonable amount of time¹. Problems that belong to *NP-Complete* on the other hand are more difficult as there exists no reliable way to generate a solution. However, if a solution is proposed, it is at least possible to recognize it as such. *NP-Hard* problems don't even have that luxury, not only is there no known

¹Incidentally, 'a reasonable amount of time' in computational complexity means 'before the universe ends', so this is not necessarily a *practical* categorization.

way to generate an answer, it isn't even clear how to test a tentative answer for correctness.

For the purposes of this paper we shall comprehensively ignore the accumulated knowledge of the past thousand years of algorithmics and approach all problems in an NP-Hard-ish fashion. That is:

- it is not known how to generate the correct solution,
- it is not known how to test a proposed solution for correctness,
- but it is possible to compare two proposed solutions and select the more correct one.

1.2 Solver categories

Solutions to problems—or rather, *methods for finding* solutions to problems—can of course also be categorized. Algorithms can be described as *greedy* or *lazy*, *stochastic* or *deterministic*, *iterative* or *recursive*, *exact* or *approximate* and a million adjectives more. These characteristics are by no means mutually exclusive or indicative of quality. Different situations call for different types of algorithms. Take for example the age-old problem of sorting a collection of values. There exist at least a dozen famous sorting algorithms², each with its own strengths and weaknesses. Which algorithm is best depends on whether one is sorting large or small collections of data, whether the data is already somewhat sorted, whether the algorithm is allowed to use a lot of memory, whether the algorithm should yield intermediate results if aborted, how unsortable data is handled and so on and so forth.

Since a number of different solver algorithms will be discussed below, a passing familiarity with some relevant algorithm categories is important. A casual definition of each will suffice in this context.

²Visit sorting-algorithms.com for a visual comparison of the seven most common ones.

Greedy algorithms are very local-minded. They make decisions based on the immediate environment rather than taking long-distance or long-term goals into account.

Deterministic algorithms follow a fixed and predictable process which tolerates no chance or randomness.

Stochastic algorithms have a random component to them and are therefore less predictable than deterministic algorithms³.

Exact algorithms are guaranteed to find the best possible solution(s) given the initial constraints.

Approximate algorithms will typically find some sort of solution without any guarantee that it is the best possible one.

Progressive algorithms compute solutions in an iterative manner, where each cycle yields—on average—a better answer than the previous one.

Adaptive algorithms can operate on a changing set of constraints and inputs. These algorithms run continuously within a dynamic environment.

Specific algorithms are designed to solve only one kind of problem. As a result they tend to be fast and reliable, but they are difficult to write. The vast majority of algorithms used on computers today fall into this category.

Generic algorithms are designed to tackle a wide variety of problems. This flexibility is accompanied by a significant drop in performance.

Open algorithms allow external entities (be they human beings or other algorithms) to participate in the solving process. Seemingly unpromising lines of inquiry can be investigated upon the request of an external actor.

³On digital computers, *all* processes are inherently deterministic, but pseudo-randomness is sufficient to classify an algorithm as stochastic.

2. PROBLEM ANALYSIS

So how exactly is a generic solver able to deal with many different problems? Doesn't that require a large amount of intelligence and understanding? The short answer is *yes*, but the key insight is that the intelligence need not reside in the solver itself. If the solver is *open*, another algorithm (which may well be classified as *specific*) can fill in the knowledge gaps. In other words, a generic solver doesn't need to know the first thing about nurbs geometry in order to find the intersection point between two curves, all it needs is a *companion algorithm* that does. By separating the 'knowing' and the 'solving' into two disjoint algorithms, they both become easier to write and easier to repurpose.

Communication between the generic and the companion algorithm can be thought of as a mixture of *twenty questions* and *hunt the thimble*. The generic algorithm confronts the companion algorithm with a tentative solution, to which the latter assigns a quality rating (cold, warmer, hot!). The generic algorithm is tasked with interpreting the clues and the companion is tasked with computing the 'temperature' or *fitness* of each proposed solution. For this exchange to work, both algorithms need to speak a common language. Happily this language is very succinct and it can be described by a single mathematical equation:

$$f(\tau) = q \quad (1)$$

$$f : \tau \rightarrow \mathbb{R} \quad (2)$$

Equation (1) defines the common language as a mathematical function, while equation (2) shows the mapping of this function, which is the mathematical way of specifying what sort of data goes in and what sort of data comes out. Terse though this notation may be, it hides a number of fairly abstract ingredients and deserves a detailed discussion.

In the above notation f represents the fitness function, which operates on τ (more on that later). The output of f is labelled q (for quality) and it

represents the fitness as a single numeric value⁴. The mapping notation (2) merely states that f has to consume data in the form of τ while it should return data in the form of a real number.

This leaves us with τ , which is a type of *tensor*. A tensor is simply a collection of variables called *elements*, which together describe all possible answers to the problem at hand, both the good and the bad ones. In fact the precise layout of τ depends on the nature of the problem, and defining this layout is the responsibility of the companion algorithm. Since digital computers are only capable of dealing with numeric data, all elements that make up τ must be numbers.

Along with a tensor definition, the companion algorithm must also specify which transformations can be applied to τ . Usually this involves nothing more complicated than a list of directions in which τ is allowed to be pushed and—for each direction—how far one can push it.

Perhaps an example is the best way to explain what a tensor in this context is. Imagine one is asked to paint a 10 cm dot on the Mona Lisa so that it least disrupts the original painting. This is basically an optimization problem of two variables; *position* and *colour*. We need to find which combination of position and colour has the highest fitness. However position and colour are not numbers and thus cannot directly be elements of τ . Luckily it is possible to describe positions and colours using numbers, which provides a way forward for defining a tensor for this particular problem:

$$\tau = \{X, Y, R, G, B\} \quad (3)$$

The position variable is represented by two numbers that encode the horizontal and vertical offset of the paint dot, while the colour variable

⁴It has to be a single number as it is vital that two fitnesses can be unambiguously compared using relational operators. If the fitness were more than a single number, then the problem can be described as *multi-variate*, and in such cases solver algorithms are employed to map out *families of solutions* that balance multiple conflicting fitness metrics. This paper will not discuss such cases.

requires three numbers, one for each of the RGB channels. A tensor element is defined by two properties; namely the *index* at which it occurs in the tensor, and also a *set of possible values* it can assume⁵. The *R*, *G* and *B* elements for example can be assigned any integer value in the range 0–255. If we assume the dots can only be positioned on the painting at one centimetre intervals, then *X* and *Y* are constrained to the integers in the ranges 5–48 and 5–72 respectively⁶. These limits make it possible to calculate the total number of different *states* τ can assume. The number of states, also called the *variability* of a tensor, can be written using angle bracket notation:

$$\langle \tau \rangle = \langle 44, 68, 256, 256, 256 \rangle \quad (4)$$

$$\langle \tau \rangle = 44 \times 68 \times 256^3 \approx 5 \times 10^{10} \quad (5)$$

There is now enough information to do some actual maths. The tensor for this problem contains five elements which are entirely independent of each other, meaning that the amount of red in the colour is not limited or affected by the amount of blue or the x-position of the dot. Five independent elements define a five-dimensional tensor. The variability of τ equals the product of the variabilities of all the elements of τ , as per equation (5). Even if a single call to $f(\tau)$ took only ten milliseconds, it would still take nearly sixteen years to iterate over all unique states of τ . This is clearly not a practical approach.

The set of all states of τ is called a *phase space* and is denoted with the symbol \mathcal{P} (a superscript integer is sometimes included to denote dimensionality). Every possible location in \mathcal{P} is identified by a specific tensor. In this sense, tensors act as coordinates. Usually the phase space can be thought of as an N-dimensional hypercube, which

is at least a reasonably fathomable concept⁷.

A note on phase space sizes. The solver algorithm mentioned above (i.e. ‘try every possible answer and remember the best one’) is called a *brute force search*. For small problems brute force may be a reasonable solution, but as the problem grows larger the time it takes to enumerate and test all possible states increases rapidly. Table 1 shows brute force runtimes for several different tensor definitions, while assuming a 10ms duration for a single measurement. The rows contain tensors with one, two, three, or four elements, while the columns contain different variabilities of the tensor elements.

$ \tau $	10^1	10^2	10^3	10^4
1	0.1s	1s	10s	2m
2	1s	2m	3h	12d
3	10s	3h	116d	317y
4	2m	12d	317y	3My

Table 1: Brute force runtimes

A tensor with three elements (6) where every element can take on one thousand unique values (7) results in a phase space size of one billion unique coordinates (8). Even when a single iteration takes only ten milliseconds, to evaluate each and every one would still take 116 days.

$$\tau = \{A, B, C\} \quad (6)$$

$$\langle \tau \rangle = \langle 10^3, 10^3, 10^3 \rangle \quad (7)$$

$$\langle \tau \rangle = 10^9 \quad (8)$$

⁵In the case of interdependent elements, additional relational properties come into play.

⁶The dimensions of the Mona Lisa are 53 cm \times 77 cm, however dots cannot be placed closer to the edge than 5cm, otherwise they would only partially cover the painting, which would be cheating.

⁷In the more complicated case of interdependent tensor elements, it could be that the variability of one element (i.e. whether it can assume 1, 14 or 3517 different states) depends on the value of another element. In these cases the dimensionality of \mathcal{P} may differ from point to point. For the remainder of this paper we will ignore variable dimensionality.

2.1 $N_{\frac{1}{2}}$ dimensional landscapes

It is not feasible to properly convey what a complete phase space of more than two dimensions looks like on paper, which is why I will adopt the space-time diagram convention of cosmology and draw the phase space as a flat, two-dimensional plane. In such a representation the individual tensors that populate the space are all sitting side by side on the plane.

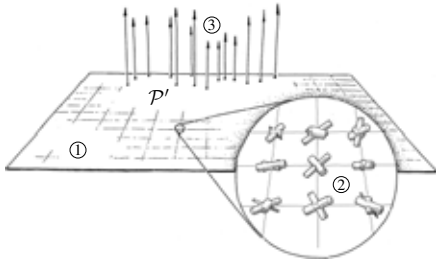


Figure 1: \mathcal{P} as the set of all tensors

In the above figure ① represents the entire phase space \mathcal{P} as collapsed to a plane⁸. The objects marked ② represent the individual tensors (visualised as cylinder permutations rather than collections of numeric values). The vertical bars marked ③ represent the fitness values of all tensors. These values are computed by the companion algorithm via $f(\tau)$.

The notion of a *fitness landscape* \mathcal{L} as the complete set of all fitness values now starts to emerge. That is, if we apply $f(\tau)$ to every τ in \mathcal{P} and thus draw all possible arrows, we get a scalar field of fitness values that pervades all of \mathcal{P} . The easiest way to imagine this geometrically is to treat the fitness value at each point in \mathcal{P} as an elevation, thus ‘pulling’ \mathcal{P} from a flat plane into a landscape.

⁸As a reminder that this is a *projection* of \mathcal{P} onto two dimensions, this paper will include a \mathcal{P}' symbol in any figure which contains such a collapsed space.

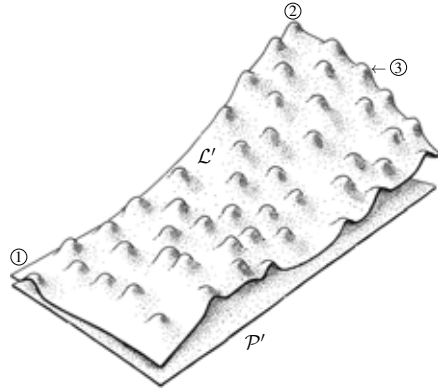


Figure 2: Fitness landscape as an extrusion of \mathcal{P}'

Figure 2 shows the fitness landscape resulting from a fitness function being applied to a phase space⁹. In the landscape one can immediately see certain characteristics of \mathcal{P} and $f(\tau)$ such as low quality solutions ①, high quality solutions ②, and local optima ③.

Since the mapping of the fitness function is defined as $\tau \rightarrow \mathbb{R}$ (see equation (2), page 3), we only get a single fitness or ‘elevation’ for every point in \mathcal{P} . Which means that planes, mountains, and valleys are all possible features of fitness landscapes, but caves, bridges and overhanging cliffs are not, as that would require more than one elevation value per coordinate. In this sense the landscape falls ever so slightly short of being truly three-dimensional. This sort of geometry is typically referred to as two-and-a-half dimensional in the machining and graphics industries¹⁰. Since all phase spaces are of integer dimensionality, it follows that all fitness landscapes are of integer dimensionality + $\frac{1}{2}$.

⁹Again, to avoid misunderstandings, fitness landscapes that are drawn as $2\frac{1}{2}$ dimensional objects are marked with an \mathcal{L}' symbol.

¹⁰Not to be confused with the partial or ‘Hausdorff’ dimensions that occur in fractal geometry.

Although it is not possible to draw landscapes whose dimensionality exceeds $2\frac{1}{2}$, it is perhaps possible to describe them verbally and get a sense of how the complexity increases with every additional dimension. A one-dimensional phase space can be represented by a line, like the x axis on a typical graph. The associated landscape must therefore also be a single curve that moves up and down as one travels from left to right along \mathcal{P}^1 , see figure 3. The best solution can be found wherever the curve has its highest peak, which is a trivial exercise for the human eye (unless there are multiple peaks with very similar elevations). One could spot such a peak after looking at a curve for no more than half a second.

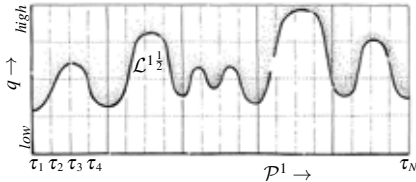


Figure 3: $f(\tau)$ for \mathcal{P}^1 .

In the two-dimensional case, \mathcal{P}^2 and $\mathcal{L}^{2\frac{1}{2}}$ actually look like the diagrams in this paper, but the geometry becomes significantly more complex with \mathcal{P}^3 and $\mathcal{L}^{3\frac{1}{2}}$, as it is not possible to imagine an extrusion at right angles to a volume. However instead of a landscape, one can think of an elevator car with cigarette smoke. The interior of the car is an analogy for \mathcal{P}^3 and the density of the smoke represents the fitness at each point in \mathcal{P}^3 . The smoke will be thicker in some places and more spread out in others. Instead of a peak in a landscape, the search is now for the cubic millimetre with the thickest smoke. This will certainly require a lot of squinting and head bobbing, if one can see it at all.

Extending this analogy to one more dimension (\mathcal{P}^4 and $\mathcal{L}^{4\frac{1}{2}}$) requires that one take into account not just the smoke at any one instant in time, but the motion of smoke particles during an entire time interval. In this case the goal is to find the

cubic millimetre with the highest density at any given second between the doors closing in the lobby and the doors opening on the fifth floor. Every additional dimension increases the size of \mathcal{P} in an exponential fashion, which is why the runtimes in table 1 increase so dramatically with every row.

3. THE FAMOUS FOUR

Up to this point the purpose of this paper has been to introduce the theory of tensors, phase spaces, fitness functions, and fitness landscapes. While it is possible to draw entire landscapes as a theoretical exercise, one should bear in mind that at the start of a new search, the actual shape of a fitness landscape is entirely unknown. The diagrams in this paper serve only as visual aides and one should not infer from them that these landscapes are at any time entirely computed. There is one thing—and one thing only—that can be done to gain information regarding the shape of the landscape, and that is to pick some tensors and see what fitness values $f(\tau)$ assigns to them.

One can think of it as standing in the landscape and having a GPS receiver but no map, while being surrounded by such thick fog that even the terrain underneath is barely visible. Given these limited tools, how does one find high-ground? This is the issue that generic solvers must deal with, and different solvers take different approaches. Like a GPS reading, a single sample may take a significant amount of time. Testing a bunch of pixels underneath a circle may take no more than 10ms, but performing a sunlight analysis on a specific window shape for every daylight hour for every day of the year could easily take seconds if not minutes. The solver must find high-ground in the landscape before the passage of the time renders the problem irrelevant.

The remainder of this paper will discuss four common generic algorithms in order of complexity while highlighting their strengths and weaknesses against the backdrop of different landscape geometries and topologies.

3.1 Divide and conquer

Divide-and-conquer (DC) is the name for an entire paradigm of algorithms, both specific and generic ones. The core idea behind DC is to break a problem into smaller and smaller pieces until the fragments are small enough to be solved directly (for example using a brute force algorithm). An implementation of DC within the context of peak finding might work as follows:

1. Define a domain \mathcal{D}_i for the search. When the algorithm begins, \mathcal{D}_i will be identical to the boundaries of \mathcal{P} .
2. Sample the tensors at the vertices v of a grid \mathcal{G}_i within \mathcal{D}_i .
3. Find the highest quad(s) within the grid.
4. Shrink \mathcal{D}_i to contain only these quads. This may involve splitting \mathcal{D}_i into one or more disjoint \mathcal{D}_{i+1} regions.
5. Repeat (1) through (4) until:
 - (a) a solution with an acceptable (predefined) fitness is found,
 - (b) or \mathcal{D}_i becomes so small as to make further subdivisions pointless.

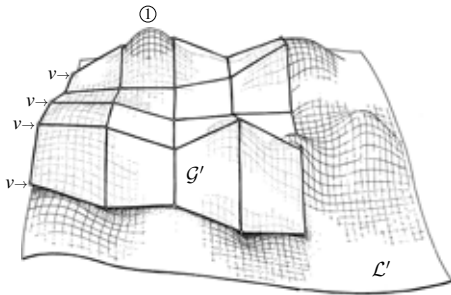


Figure 4: Sampling grid on \mathcal{L}

Such an algorithm is easy to implement, but it does rely on a number of assumptions regarding \mathcal{L} that will not be true for all cases. First of

all the algorithm treats the sampling grid \mathcal{G} as a proxy for \mathcal{L} at every iteration. If \mathcal{L} contains relevant details that are smaller than the spacing between adjacent samples, they may well go overlooked. Narrow peaks with a small footprint are particularly likely to go unnoticed, as ① in figure 4 shows. To rephrase that in a way that points the accusing finger at the algorithm instead of the landscape; DC assumes that \mathcal{L} is uninteresting over short distances and can thus be reliably approximated with a low density grid.

The extent of a peak, usually called its *basin of attraction*, can be described as the area from which one can reach the peak while only walking straight uphill. The height of a peak and the extent of its basin are completely unrelated properties, and quite often peaks representing low quality solutions have larger basins than peaks representing high quality solutions.

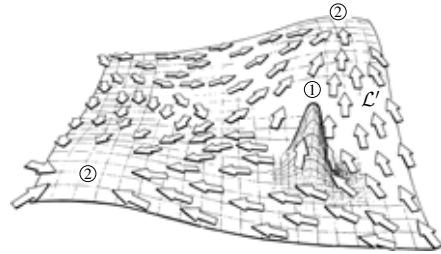


Figure 5: Basins of attraction

The above figure shows a landscape with three peaks, where ① represents a high quality solution, while the remaining ones marked ② represent low quality solutions. The arrows show the steepest way up from a sampling of points across \mathcal{L}' . As is now visible, the vast majority of arrows lead to low quality peaks.

Another problem with DC is that it does not scale well to higher dimensions. A sampling grid accuracy needs to be picked for every iteration, and this accuracy needs to be tight enough to not miss significant landscape features. The number of actual fitness samples that needs to be taken

for a single DC iteration¹¹ is defined as the product of the number of samples for every tensor element. For a sampling density of ten steps per tensor element in \mathcal{P}^4 , a single iteration requires $10^4 = 10,000$ samples. At half a minute per sample for something as computationally intense as for example a daylight analysis, the first DC iteration will have a runtime of about three and a half days.

Yet for low-dimensional problems DC is often a good approach and it is certainly one of the more popular search algorithms out there. A straight-forward implementation of DC can be categorized as greedy, deterministic, approximate, and progressive (see page 2 for an explanation of these categories). Since the algorithm quickly discards large areas in the landscape as it shrinks the search domain, it is not adaptive.

3.2 Hill climbing

While divide-and-conquer at least attempts to find the best solution, a hill climbing algorithm is designed to only take local conditions into consideration. Starting from an arbitrary location on the landscape it will try and walk uphill as fast it can. It achieves this by sampling the tensors adjacent to the current position to see which one is the fittest and then moving in that direction until it finds a peak, or until something bad happens. A simple implementation of a hill climber might work as follows:

1. From a location τ_i , sample adjacent tensors in all meaningful directions. Remember τ_j as the neighbour with the highest fitness.
2. Define a travelling vector $\vec{v} = \tau_j - \tau_i$, and multiply by a scalar S representing the step size.
3. Take a step in \mathcal{P} , defined as $\tau_{i+1} = \tau_i + \vec{v}$.
4. Evaluate the fitness q_{i+1} at τ_{i+1} .

¹¹At least assuming that a search domain does not split into multiple subdomains, in which case more samples are required still.

5. Repeat (3) and (4) until:

- (a) q_{i+1} is less than q_i ,
- (b) or until a boundary of \mathcal{P} is reached.

6. Repeat (1) through (5) until:

- (a) a solution with an acceptable (predefined) fitness is found,
- (b) or there are no adjacent tensors with higher fitness than the current one.

To put that in plain English; pick a direction that seems to go uphill the steepest, then keep walking in that direction until you start walking downhill, or until you can walk no further. Then pick a new best direction and start walking again.

Hill climbers are greedy algorithms that tend to find the peak whose basin of attraction they start in. This behaviour isn't necessarily a drawback, but it does mean that using hill climbers to find the best possible solution(s) is only possible if \mathcal{L} doesn't have too many local optima. Bumpy landscapes such as shown in figure 2 are unsuitable for a hill climbing approach. In extreme cases \mathcal{L} can start to exhibit fractal properties (bumps upon bumps upon bumps). Since most algorithms depend on at least a small level of continuity in \mathcal{L} , such landscapes often defy navigation algorithms, but hill climbers in particular are vulnerable to such geometry.

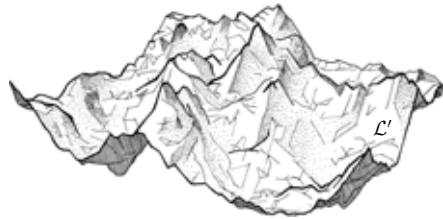


Figure 6: Landscape with fractal properties

Hill climbers are significantly more difficult to implement than divide-and-conquer algorithms, mostly because there is a lot of fine-tuning

involved. Although in theory adjacent tensors can be evaluated to get the local *slope* or *gradient* of \mathcal{L} , it is sometimes better to pick tensors that are not directly adjacent but a bit further away. Short distance sampling is liable to suffer from tiny amounts of numeric noise that are inherent in $f(\tau)$, but how far away should a sample be taken? It depends on the nature of \mathcal{L} , which the algorithm is not supposed to know anything about.

Another fine-tuning problem arises with the selection of a step size S . The best sampling radius for the gradient approximation is unlikely to be the same number as the ideal distance for subsequent steps. But what is the ideal value for S ? Again, it depends on the geometry of \mathcal{L} .

It is also reasonable to expect S to decrease over the course of the search. While running towards a far away peak, it makes sense to take long strides in order to get there quicker. But near a peak big leaps will probably just overstep the summit and make things worse. But how sharply should S decline over time? Nobody knows.

In addition to the problems mentioned above, which are at least in principle solvable, hill climbers suffer from poor scalability to higher dimensional spaces. It is computationally cheap to walk across a landscape, but as the dimensionality of \mathcal{P} increases it becomes more and more expensive to determine the local gradient¹². One would need at least two samples in opposite directions for every principle axis in \mathcal{P} ¹³.

When discontinuities in \mathcal{L} are rare (even if they are large), they should not interfere too much with a hill climbing approach, as there are still plenty of connected areas where a sequence of progressive steps can be taken.

¹²Measuring the local gradient is only expensive if it has to be *inferred* by sampling the immediate surroundings. If a derivative fitness function $f'(\tau)$ is known then the search would scale very well to higher dimensions. However, a lot of fitness landscapes are not readily differentiable.

¹³This is true for tensors on the interior of \mathcal{P} only. Locations along the boundaries and edges of \mathcal{P} require fewer samples.

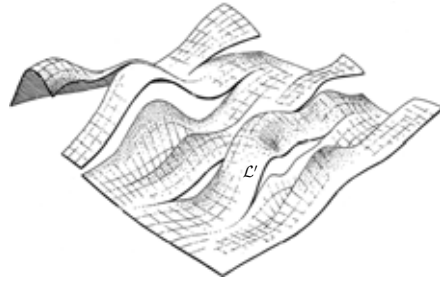


Figure 7: Landscape with few discontinuities

Hill climbing algorithms can be categorized as greedy, deterministic, exact (though only in a *local* optimum sort of way), and progressive. It is possible to introduce a certain amount of randomness and turn the categorization away from the deterministic and towards the stochastic. Instead of sampling tensors parallel to phase space axes, a randomized sampling could be employed. One benefit of such a change is that it may reduce the number of turns required, especially in the case of ridges that exist at angles to tensor axes in \mathcal{P} .

3.3 Simulated Annealing

The previous two algorithms may not be simple in their details, but they mimic the way humans would solve problems and it is therefore easy to see how they might work. *Simulated annealing* is not like that and it appears to follow some very unusual and suspect rules. ‘Annealing’ is a term that comes from a treatment process in metallurgy which is closely related to crystallization.

The atoms in most metals and alloys like to sit on a regular lattice called a *crystal*. Atoms which are aligned with the lattice exist in a *minimum energy state*. However in the real world, a sample of metal will contain dislocations and cracks and adjacent regions of differently oriented atoms. These departures from a perfect crystal are called *defects*. By heating up the sample, one can weaken the bonds between atoms which enables them to randomly jump around a bit until they find a lower energy state than they had before. As the sample cools down, atoms

will tend to settle in these new states and the sample will end up with larger regions of pristinely crystallized atoms and fewer defects.

The analogy with *simulated* annealing as a problem solving approach is somewhat weak, but the mathematics are very similar. A simulated annealing solver takes the thermodynamic equations that describe metallic annealing and applies them to tensor movement. As a result the algorithm is highly stochastic, fairly progressive, somewhere in between exact and approximate (depending on the complexity of \mathcal{L}), and slowly transitions from global to greedy during the progression of each search.

A basic simulated annealing solver is actually remarkably easy to implement and could work as follows:

1. Define a temperature \mathcal{K} . At the start of the search \mathcal{K} should be high.
2. From the location of τ_i , randomly pick a τ_{i+1} somewhere within a maximum radius d , where d is covariant with \mathcal{K} . The higher the temperature, the larger the distance.
3. Evaluate the fitness q_{i+1} at τ_{i+1} .
 - (a) If $q_{i+1} > q_i$ then move to τ_{i+1} .
 - (b) If $q_{i+1} < q_i$, then maybe *still* move to τ_{i+1} , depending on the difference in fitness (Δq) between q_i and q_{i+1} and \mathcal{K} . The higher the temperature and the smaller the difference in fitness, the more likely a move to a less fit tensor will be accepted.
4. Redefine \mathcal{K} as a percentage of its current value, simulating a cooling environment.
5. Repeat (2) to (4) until:
 - (a) a solution with an acceptable (predefined) fitness is found,
 - (b) or until \mathcal{K} drops below a predefined limit.

The surprising feature of simulated annealing is of course that it will sometimes deliberately move to a worse solution. The benefit of being willing to make things worse is that it sometimes allows one to break out of a local optimum and find higher ground that is more than one move away. Since \mathcal{K} is reduced during every iteration though, the algorithm becomes less and less willing over time to accept a move to a lower quality tensor. One can imagine simulated annealing as two solvers wrapped into a single package. At first the search includes all of \mathcal{P} and the algorithm is clearly global minded. It is during this phase that high ground is sought. Then as \mathcal{K} gets lower and lower the algorithm instead starts behaving like a greedy stochastic sampler, only willing to climb the local peak.

There are very few drawbacks to simulated annealing and the things that do plague it tend to wreak havoc no matter what generic solver you apply. In particular high peaks with small basins of attraction as shown in figure 5 can easily be missed by the random nature of the sampler. On the other hand simulated annealing can deal very well with discontinuous and fractal terrain.

The problem I would like to discuss in this context is not associated with simulated annealing solvers but rather with poor implementations of $f(\tau)$. It is entirely possible to write a fitness function that yields identical fitness values for adjacent tensors. This will result in flat areas or *plateaus* in \mathcal{L} as shown in figure 8.

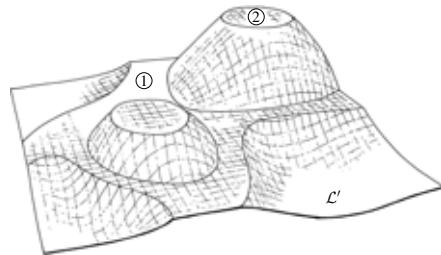


Figure 8: Geometry of under-constrainedness

Plateaus in fitness landscapes are a manifestation of *under-constrained* fitness functions. Let us repurpose the painting-a-dot example discussed on page 3, but instead of the *Mona Lisa*, this time the painting in question is *Who's Afraid of Red, Yellow and Blue*. In this new case, a lot of red dots can clearly be painted which are all equally fit in the sense that they do not change the appearance of the painting in any way. In other words, the 'peak' in the fitness landscape defined by this \mathcal{P} and this $f(\tau)$ is more like a *mesa* than a summit.

Plateaus in the low or medium ground of a fitness landscape may well retard the progress of many solvers as it becomes very difficult to decide in which direction it is smart to move. But plateaus in the high ground of a landscape, specifically if they are the highest features, may result in different yet equally valid solutions. This is rarely what people are after but luckily it is often possible to add additional terms to $f(\tau)$ which introduce a gradient to \mathcal{L} so that perfectly horizontal plateaus become ever so slightly tilted or curved.

Note that a plateau is defined as two or more adjacent tensors with identical fitness. This means that the dimensionality of plateaus can be any number up to and including the dimensionality of \mathcal{P} as shown in figure 9.

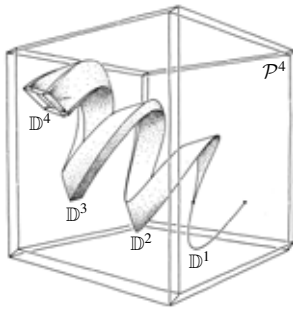


Figure 9: Plateaus of different dimensionality in a 4-dimensional phase space.

3.4 Simulated Evolution

Evolutionary solvers and *genetic algorithms* have long been popular amongst the computationally minded and there are many different ways to implement them. The ideas behind such algorithms are derived from biological evolution and its two main components of mutation and selection. Evolution is nature's answer to systematic problem solving, but apart from being very successful, it is also rather slow. In nature that tends not to matter too much because of the availability of massive parallel processing (every atom simulates itself in real-time), but when simulating the process digitally, the performance can become prohibitively slow.

The core idea of both biological and simulated evolution is the *heritability of traits*. Biologically, traits are phenotypic expressions of allele-complexes and the fitness of a specific complex is measured by its frequency in the gene-pool over time. Things are much simpler in a simulated evolution environment as the fitness is not an intrinsic property that is manifested statistically over several generations, but rather it is defined directly by $f(\tau)$. The analogy between the biological terms and mathematical symbols can be summarised as follows:

- Population $\longrightarrow \hat{\mathcal{X}}$
- Generation $\longrightarrow \mathcal{X}_i$
- Genome $\longrightarrow \tau$ (tensor)
- Gene $\longrightarrow \tau_i$ (tensor element)
- Allele $\longrightarrow |\tau_i|$ (element value)

A typical evolutionary solver is an iterative algorithm which maintains a *population* $\hat{\mathcal{X}}$ of individuals and selectively culls and recombines them to form successive generations \mathcal{X}_i , \mathcal{X}_{i+1} , \mathcal{X}_{i+2} , etc. The point is to only use the fittest individuals to form the next generation in order to select from and expand upon the best genes from a randomly generated gene-pool. The founder generation \mathcal{X}_1 can either be distributed randomly across all of \mathcal{P} , or it could be limited to a smaller

patch in order to explore local optima. Not all generations need have the same number of individuals. Especially \mathcal{X}_1 may benefit from being much larger than subsequent generations, as a dense sampling of \mathcal{L} will reduce the chances that solutions with small basins of attraction will go unnoticed. The entire process can be described as:

1. Populate \mathcal{P} with a founder generation \mathcal{X}_1 of randomly picked genomes.
2. Evaluate each genome in \mathcal{X}_i using $f(\tau)$ and create a hierarchy \mathcal{H}_i based on their fitness distribution.
3. Combine genomes from \mathcal{X}_i based on \mathcal{H}_i to create offspring genomes for \mathcal{X}_{i+1} .
4. Repeat (2) and (3) until:
 - (a) a genome with an acceptable (predefined) fitness is found,
 - (b) or \mathcal{X}_i fails to significantly improve upon \mathcal{X}_{i-1} , \mathcal{X}_{i-2} , ...

But succinct as the basic approach may be, there are a lot of details left unsaid and a lot of decisions left unmade. Item (3) especially hides a lot of complexity. The purpose of iteration in this algorithm is to evolve a population of tensors that occupy high peaks in \mathcal{L} . For this to work, the fittest tensors in generation \mathcal{X}_i have to combine to produce even fitter tensors in generation \mathcal{X}_{i+1} , but a lot can go wrong in this process.

Offspring from nearly coincident parents will not contribute significantly to the representation of \mathcal{L} that the algorithm builds over time, since that particular region has already been sampled. Measuring the fitness of highly related tensors very quickly suffers from diminishing returns. On the other hand, parents which are too far apart may well belong to different *sub-species*, each in the process of climbing a different peak. Since offspring will most likely fall somewhere in between the parents in \mathcal{P} , they are likely to end

up in a valley, as valleys tend to separate peaks in a landscape. These two mating extremes are called *incestuous* and *zoophilic* respectively and they are as detrimental in nature as they are in simulations.

Ideally one would select parents that are positioned on opposite sides of the same peak. But even in the ideal case there are still more decisions to be made regarding the combination of two parent genomes to produce offspring. This process is called *coalescence* and several approaches exist.

Coalescence where some elements are copied intact from τ_1 and others from τ_2 most resembles the biological process of fertilization where gametes combine to form a new genome. This kind of merging only works well when τ contains many elements:

$$\begin{array}{c} \tau_1 \{ \underline{3}, 8, 1, 17 \} \\ \tau_2 \{ \underline{5}, \underline{2}, \underline{9}, \underline{25} \} \\ \downarrow \\ \tau_3 \{ 3, 8, 9, 25 \} \end{array}$$

Coalescence where alleles are interpolated between τ_1 and τ_2 treats the genes as smoothly varying properties. In this sense the tensor elements do not so much resemble different alleles as entire allele-complexes. The interpolation need not always be halfway, parental genomes can be weighted either randomly, or based on Δq :

$$\begin{array}{c} \tau_1 \{ \underline{3}, \underline{8}, \underline{1}, \underline{17} \} \\ \tau_2 \{ \underline{5}, \underline{2}, \underline{9}, \underline{25} \} \\ \downarrow \\ \tau_3 \{ 4, 5, 5, 21 \} \end{array}$$

Finally, there's no reason why (in a simulation) offspring requires two parents. One could instead take all τ s in \mathcal{X}_i into consideration during coalescence, or perhaps simulate asexual reproduction.

In addition to breeding selection, partner selection and coalescence algorithms, an evolutionary solver should also provide ways for introducing mutations into a genome. Again, there are many ways to approach this problem, whether

one wants to treat genes as independent or covariant, whether mutations are associated with gene blending due to coalescence, whether mutations are dependent on the uniformity of all tensors in the (sub)species, whether mutations are in some way related to \mathcal{H}_i or even the recent history of \mathcal{H} over several generations, and so on and so forth.

Like all peak finding algorithms, evolutionary solvers perform better on certain landscape geometries. As with all solvers, peaks with small basins of attraction are easily missed and plateaus are crossed slowly. Unlike many other algorithms however, evolutionary solvers are capable of exploring multiple peaks simultaneously by populating each peak with its own sub-species. As a result evolutionary solvers are decidedly non-greedy, reasonably adaptive (especially if \mathcal{L} changes *slowly* over time), particularly open, and very stochastic.

Unlike hill climbers and simulated annealing, evolutionary solvers don't really 'walk' across the landscape in search of peaks. Instead they act more like expanding and thinning clouds of tensor particles. The velocity of the cloud front is typically quite low and it is easily deflected by obstructions such as plateaus.

While plateaus, narrow peaks and fractal terrain are all geometric properties, *over-constrainedness* causes *topological* defects in fitness landscapes. When a fitness function is over-constrained, it becomes impossible to evaluate certain states of τ , and as a result \mathcal{L} will contain gaps. Depending on the size and connectivity of these gaps, and depending on the solver in question, this may or may not be a problem. Even solvers that seem to move across \mathcal{L} in a continuous fashion do not actually travel from τ_i to τ_{i+1} in a smooth fashion, they jump instantaneously from one to the other. If a piece of \mathcal{L} is missing in between τ_i and τ_{i+1} nobody will be the wiser. But as gaps grow bigger, so do the chances of stepping in one.

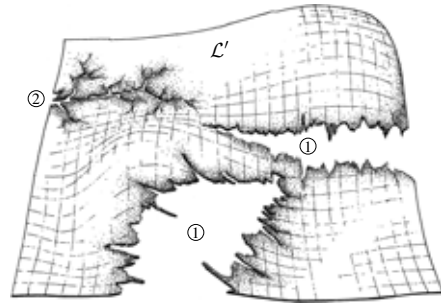


Figure 10: Geometry of over-constrainedness

There are two important gap metrics in dealing with over-constrained landscapes: *area* and *circumference*. Gaps covering large portions of \mathcal{L} will delay or even halt a solver as it is unable to cross over to the other side. Small gaps on the other hand may go completely unnoticed as the solver never attempts to sample an over-constrained tensor. The lower limit for the circumference of a gap is at least equal to the circumference of a circle of equal area¹⁴, but there is no upper limit. The boundary of a gap can exhibit fractal properties which can trap algorithms such as hill climbers and evolutionary solvers into ever narrowing tendrils of \mathcal{L} , see ① in figure 10. This sort of topology acts as an over-abundance of local optima. However if the area is small—even in the case of fractal boundaries—then the gap can go unnoticed as it will tend to fall between successive samplings of \mathcal{L} , see ② in figure 10. Simulated annealing in particular is adept at dealing with over-constrained problems, as it often traverses large distances and can thus easily span gaps that cover significant percentages of \mathcal{L} .

¹⁴Technically the minimal circumference (as measured in \mathcal{P} , not \mathcal{L}) is equal to the area of a hyper-sphere with the same volume and dimensionality as the gap.

4. CONCLUSIONS

This paper discussed the fundamentals of generic solvers. Although a detailed understanding of such algorithms is not a prerequisite for their application, potential users should at least familiarise themselves with the notions of phase-spaces, fitness functions and fitness landscapes, as defining these aspects is the responsibility of the user. It is important to realise that generic solvers are stochastic rather than analytic processes, and that they may take a prohibitively long time to run, depending on the specifics of the problem at hand.

ACKNOWLEDGEMENTS

This paper was written in L^AT_EX 2_ε, proofread by K., and funded by Robert McNeel & Associates. Thanks to all.

ABOUT THE AUTHOR

David Rutten is a graduate of the faculty of Architecture and Urbanism at TUDelft. Since 2006 he has been employed by the *Andrew leBihan partnership* in *Turku, Finland*, and, more recently, by *Robert McNeel & Associates* in *Seattle, WA* where his main task is the continued development of and support for the Grasshopper[®] plug-in for Rhinoceros 3D[®]. Grasshopper is a visual programming environment aimed at those who wish to partake in computational design. David works from home in an undisclosed location in Tyrol.

NAVIGATION FUNCTIONS OF PICTORIAL SCHEMATICS DURING THE EDO PERIOD WITH EMPHASIS ON TRAVEL MAPS

Toshimasa KONISHI and Shinya. KANEKO
Sapporo City University, Japan

ABSTRACT: The results of studies focused on centers of sight and lines of vision in pictorial drawings during Japan's Edo Period have been announced at meetings of the ISGG between the years 2002 and 2008. The characteristics of Japanese pictorial drawings are such that a single drawing contains a number of centers of sight and lines of vision. Discussions have been held as to how these are peculiar to such drawings. The reasons as to why such centers of sight and lines of vision are repeatedly incorporated have been made clear.

At the meeting of the society's Montreal Chapter, held in 2012, a comparative analysis was made between Edo pictorial drawings and today's car navigation systems.

In both cases, the lines of vision and centers of sight are made multiplex, giving them navigation systems that are easy to understand.

At this time the focus will be on travel maps of the Edo Period, which had the objective of guiding the way. The functions of such maps have been compared with modern car navigation systems, which provide guidance for the drivers.

The results of the studies revealed that, in addition to use for travel, recreation and pilgrimages, one could place them on a desktop and obtain a bird's eye view of an extensive area. In other words, they can provide different expressions. When one makes a comparison between car navigation systems and pictorial drawings, for use in travel and to move from one location to another, the method conveying an exact location and how to guide people to their desired destination, including identification as to what exists between one's present location and where one is going, can be described in a very similar fashion. Even though there are significant differences between what appears on screens and on maps, the screen being sometimes fixed and sometimes moving, there are definite similarities.

Keywords: Pictorial Map, The Edo Period, Car Navigation Systems, Center of sight, Line of vision.

1. INTRODUCTION

At the 2012 conference of the ISGG, held in Montreal, a comparative study/analysis was made between today's car navigation systems and the pictorial maps of the Edo period, indicating easy- to- understand navigation functions through multiplexing the centers of sight and lines of vision. At this time, emphasis was placed on travel maps of the Edo period, made with

the purpose of guidance along the road, including leading people in the right direction. These functions and roles were analyzed; and comparisons were made between them and modern car navigation systems, the modern guidance method.

When one makes a comparison between car navigation systems and pictorial drawings, for use in travel and to move from one location to another, although there were significant differences in that one

was mobile and the other static, there were also great similarities in that both showed one's exact location and how to get to the desired destination, as well as what lay along the way.

2. Pictorial travel maps of the Edo Period

In addition to use for travel, recreation and pilgrimages, one could place so-called travel maps on a desktop and obtain a bird's-eye view of an extensive area. In other words, they can provide different expressions. Needless to say, car navigation is something that one must have, even in cities. A comparison is thus in order.



Figure 1: Map of Yonezawa Domain, Dewa Province, 1831, 590 x 784

As for the type that provides a desktop view of a wide-ranging area, there are many similarities with Western pictorial maps, with hardly any differences. However, in the case of large Japanese pictorial maps from the Edo period, the map is placed on the floor, rather than being hung on the wall, in order to be seen from a circumferential point of view. Writing is included along each border, in order to make it easier

to read the map from that side (Figure 1). A form of this type of map, containing writing only, can often be seen.

No one knows of a case in which lines of vision toward various subjects in pictorial maps have been altered in order to facilitate easier viewing from all four edges of the map. The perspectives of the objects, such as dwellings, trees, mountains, bridges, etc., described in pictorial maps of the Edo period illustrate the scene as viewed from above the road at each point. Thus, the objects shown in pictorial maps are illustrated in such a way that the person enters into the map itself and then looks around at his surroundings. To be more precise, the object at each point may have to be illustrated in accordance with an opened-up view. It thus differs from a navigation system which involves motion because the latter cannot be present in a map. So, it is illustrated in such a way that the lines of vision are described in a parallel sense, while centers of sight, unfolding on both sides of the road, lead in the directions in which the roads are headed. Although similar to Figure 2, the difference is that the centers of sight and lines of vision focus on both sides of the road.



Figure 2: Highroad Map, Tokaido Bunken Ezu, A part of Odawara-Fuchu, with the Orientation of each position, 1690, 269 x 6300, The lines of vision move along the road in parallel sense.



Figure 3: Map of Komaki and Nagakute Batte Owari Province, mid-late Edo, 555 x 793



Figure 4: Map of Atsumi Hot Spring, bakumatsu, 598 x 400

Thus in pictorial maps containing writing, in which mountains, structures, etc. are not described in concrete terms, the writings in the maps are positioned in such a way as to make reading easier from the periphery. However, the structures, mountains, trees, etc. are not necessarily arranged in such a way that they can be easily viewed. One gains the impression that, either the perspective is integrated in one direction or it is drawn in such a way that the viewer enters into the map itself.

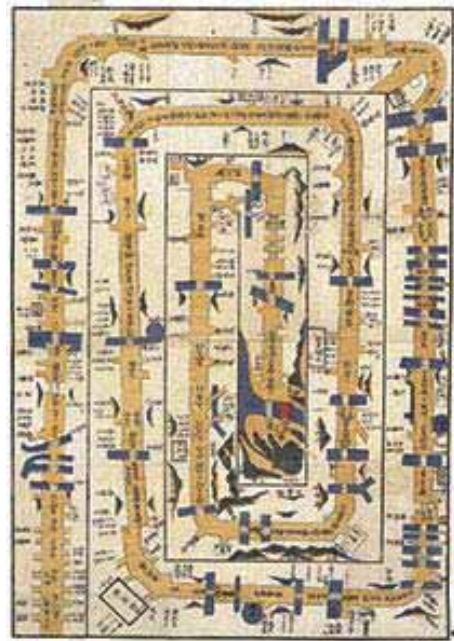


Figure 5: Highroad Map, Map of the Nikko Highroad, bakumatsu, 320 x 445

The author recently found a drawing (Figure 6 and 7) by Norimitsu Kokubo, which Mr. Kokubo drew in such a way that both sides of the road unfold, so that the viewer enters into the picture. In terms of precision, Mr. Kokubo's drawing is far superior to pictorial maps of the Edo period; but, in principle there are a number of similarities, to the extent that centers of sight and lines of vision have been moved, to focus on buildings on both sides of the road. It is also perhaps interesting that an explanation was given in writing.



Figure 6: Norimitsu Kokubo
Restoring Fukushima Prefecture East Japan (a Part),
2011, Paper, pencil, color pencil, ball point pen, 502 x
620



Figure :7 Norimitsu Kokubo (above right)
Future City under Development (a part), 2011, Paper,
pencil, color pencil, ball point pen, 1204 x 1595,

3. Pictorial maps of the Edo period and car navigation systems

Thought can be given to the effect of pictorial maps and navigation systems, anticipating their roles as means of guidance while on trips. When walking, one can stop and look around. A map can be unfolded, exposing only the part that you need. That is equivalent to a car navigation system's changing scenery.

In the initial stages of car navigation systems, one simply grasped where one was going by noting the direction of the arrow on the map. However, over time, it in some sense gradually became close to the Edo period pictorial maps, adding surrounding scenes such as the characteristics of the road, including the number of lanes, the presence or absence of sidewalks, and lines of trees. (Figure 10)



Figure 8: Screen of Car Navigation, In the initial stage, one simply grasped where one was going by noting the direction of the arrow on the map.



Figure 9: Car navigation facilitates rotation of screen, depending upon the direction in which one is going.



LARGER IMAGE



Figure 10: Some navigational screens use a birds-eye rendering of constant forward movement. Also, the scene ahead may be depicted from the driver's perspective.

Also, car navigation facilitates rotation of the screen, depending upon the direction in which one is going. It is always set in such a way that the direction you are going is forward, whereas in the case of pictorial maps one can select the direction in which to go. In doing so, one can refer to a mountain ahead or to the presence or absence of structures.(Figure 9 and 10)

As was stated at the previous meeting in Montreal, some navigational screens use a birds-eye rendering of constant forward movement. Also, the scene ahead may be depicted from the driver's perspective. A sense of a rolling viewpoint, such as in a movie, is not present in pictorial maps; but the important point is that one places oneself within the car navigation scene, similar to what one does with an Edo period pictorial map.



Figure 11: The scene ahead may be depicted from the driver's perspective, such as in a movie.

Table 1 is a summary of comparisons between pictorial maps and car navigation systems. One can compare the differences and the similarities. Other characteristics of car navigation include:

- Convenience, and accuracy in direction
- Ease of cognizance while the car is moving
- Impact of the car navigation system developed in Japan

- Greater emphasis placed on arriving at the destination, rather than locating where one is
- Ability to obtain other information, such as traffic congestion and blocked entry due to construction work

Although car navigation is convenient, there are troublesome aspects. These include:

- When car navigation stops working while driving, what should be done? In the case of pictorial maps one is aware of approximately where one is, so one can proceed, whereas if the convenient black box does not function then it becomes problematical.
- Despite that, one becomes highly dependent on car navigation, so that it is felt to be indispensable.

Table 1: A summary of comparison between pictorial maps and car navigation systems.

	Pictorial Map Edo	Pictorial Map Europe	Art Route	Car Navigation
1. Movement of Center of Sight	○	×	○	○
2. Movement of Line of Vision	○	×	○	○
3. Movement of Screen	×	×	×	○
4. Expansion of Screen	△	△	△	○
5. Perspective Depiction	○	○	○	○
6. Movement of View Center	×	×	×	○
7. Concrete Expression	○	○	○	○
8. Color	○	○	○	○
9. Change by Time	×	×	×	○
10. Explanation by Letters	○	○	○	○
11. Exaggeration	○	○	○	○
12. One's own Location	△	△	△	○
13. Planned Course	△	△	△	○
14. Programed Course	△	△	×	○
15. Partly Expansion	△	△	△	○
16. Depiction of Land Mark	○	○	×	○
17. The necessary time	△	△	×	○
18. Minimum Time	×	×	×	○
19. Maximum Time	×	×	×	○
20. Function for Guide	○	△	×	○
21. GPS.	×	×	×	○

4. Conclusion

There are interesting questions. These are : What did not exist in Western pictorial maps is now adopted and used in car navigation systems in the West. A possible reason for that is that, during the rapid movement of an automobile, decisions can be made easily. Conversely, why in the past were Japanese able to express spatial impressions in pictorial maps? The reason perhaps lies in the fact that they had a good grasp of spatial characteristics.

Future directions: There is insufficient information concerning the utility of portable, foldable old Western maps. This is a subject which will require further study on the author's part.

REFERENCES

- [1] Toshimasa Konishi, Japanese perception of space stemming from changes in centers of sight and the lines of sight in Japanese pictorial maps during the Edo period, Proceeding 12th ICGG 2006, Salvador

- [2] The committee for editing historical drawings of cities, Historical drawings of cities. Shokokusha, Spt. 1999, ISBN4-395 -00489-X.
- [3] Morrison H. Heckscher, Creating Central Park. The metropolitan museum of art, New York, Winter, 2008.
- [4] Toshimasa Konishi, Comparisons concerning perception of space in a variety of countries throughout the world, based on centers of sight and lines of vision within pictorial drawing map, proceeding 14th ICGG 2010, Dresden
- [5] Kazumasa Yamashita, Japanese Maps of the Edo Period, kashiwashobo, Oct.1998, ISBN4-7601-1670-2.
- [6] Teiji Yoshimura, Space structure of Japan, kashima shuppankai, 1982, ISBN4-306 -05173-0
- [7] Toshimasa Konishi, Lee Youngran, A comparison between pictorial maps from the Edo period, based on center of sight and line of vision, as they relate to spatial awareness in global navigation systems. Proceeding 15th ICGG in Montreal.

ABOUT THE AUTHORS

Toshimasa Konishi, Ph.D., is a Professor of Sapporo City University.
t-524@nifty.com

Shinya Kaneko, Ph.D., is Research associate of Sapporo City University.

A NEW CONDITION OF PERPENDICULARITY IN ISOMETRICAL CAVALIER PROJECTION

Vito CARDONE¹ and Barbara MESSINA²

Department of Civil Engineering, University of Salerno, Italy

ABSTRACT: The paper addresses the issue of perpendicularity in axonometric projection, almost ignored by most of the Descriptive Geometry books for students in engineering and architecture. This problem is usually approached in an operational way, identifying lines and planes, perpendicular to each other, as a solution to metric problems, but outside of an accomplished theoretical treatment, which defines the necessary and sufficient conditions of perpendicularity. The more in-depth texts, at most, provide some general mention of that problem, indicating the high road in the application the polarity.

On the basis of a student's intuition, our group has developed an unusual and interesting formulation of the perpendicularity condition, between straight line and plane, in isometric cavalier projection, precisely in "military axonometric projection". This condition completely disregards polar and anti-polar, rebatments of planes and switching in Monge method. Its geometrical proof, moreover, is very simple and the identification of the perpendicular to a generic plane requires only three steps.

The formulated perpendicularity condition, moreover, is valid for any type of isometric cavalier projection; or rather when the projection plane coinciding with one of the reference's planes, and the projector is an oblique direction, inclined by 45° to the projection plane.

Given the traditional difficulty of directly expressing the perpendicularity condition in axonometric projection, it is a very important result, not only because it allows simpler and faster graphical procedures, but mainly because it fills a real emptiness of the isometric model, offered in the courses for architects and engineers, which then appeared unfinished as completely independent model.

Keywords: descriptive geometry, isometric model, cavalier projection, perpendicularity.

1. STATE OF THE ART ²

The problem of perpendicularity in axonometric projection has always been a complex issue, almost ignored by most of the Descriptive Geometry books for students in engineering and architecture.

With reference to the most significant treaties, products between the XIX century and the beginning of the last century, there is a certain difficulty in arriving at a unique and generalized solution to the problem.

Mostly, indeed, the authors do not address this matter, preferring to dwell on the representation of surfaces, even very complex, for which, however, is not necessary the use of

perpendicularity. No mention of the issue is found for example in the essay of Jean-Nicolas-Pierre Hachette [6] or in that by Jules-Antoine-René Maillard de la Gournerie [7] leading members of the French school. As well as, no reference to the problem is in the work of the Spanish engineer Eduardo Torroja [14].

In other cases, however, the perpendicularity condition is present in the method's theoretical treatment, but with reference to the orthogonal axonometric projection. Especially engineers and mathematicians – among which is mentioned, as an example, the American William H. Roever, engineer and university professor of mathematics at Washington University – formulate a solution to the problem, following very quick and simple

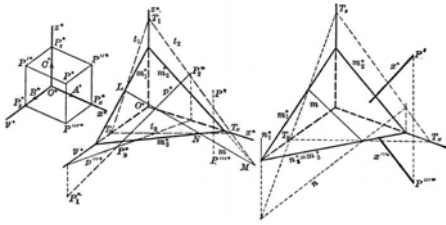


Figure 01. Straight line perpendicular to a plane in orthogonal axonometric (taken from W.H. Roever, 1918)

procedures to perform. As well summarized in 1918 by Roever «by the axonometric method, as well as by the Mongean, all the problems of space can be solved, whether these be problems of geometry of position, perpendicularity problems or metrical problems» [10].

Interestingly, although in rare situations, the problem is also dealt in cavalier projections. This is shown, for example, by the treaty of the Spanish Enrique Valenzuela [15] which, if on one hand formulates a substantially similar definition to that already mentioned for the orthogonal axonometric, on the other hand, in the chapter specifically dedicated to the cavalier projections, even focuses on the perpendicularity conditions and on the representation of a right pyramid, with base on a generic plane.

In this case – as well as in other contemporary examples that treat the cavalier projections, although not directly aimed at the perpendicularity problem – the solid's representation is performed following a hybrid process, partly synthetic and partly analytic.

That is, after fixing a reduction factor for each of the three axonometric axes, combining graphical steps to numerical calculations, are identified the associated directions to the two straight lines that define the perimeter of the solid's base, as well as the perpendicular's direction to the generic plane to which the base belongs.

In Italy, this problem is usually approached in an operational way, identifying lines and planes, perpendicular to each other, as a solution to metric problems – in the last analysis, it is in the field

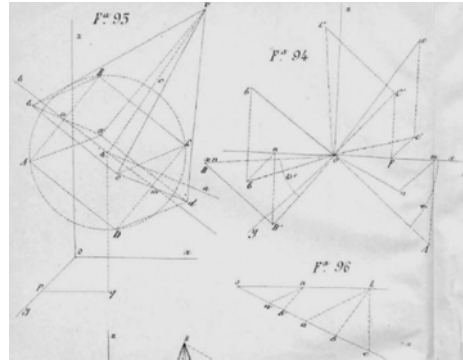


Figure 02. Representation of a right pyramid with the base on a generic plane

of angular measurements –, but outside of an accomplished theoretical treatment, which defines the necessary and sufficient conditions of perpendicularity. The more in-depth texts – such as those Luigi Campedelli [2] and Ugo Saccardi [11] – at most, provide some general mention of that problem, indicating the high road in the application the polarity.

Is precisely the important Florentine school of Descriptive Geometry, which was set up from them, in the second half of the last century addressed this topic with a certain continuity, providing some interesting contributions. For example, Aterino Aterini formulated the perpendicularity's conditions in axonometric projection, both oblique (including the cavalier projection) both orthogonal, using the anti-polarity [1].

Saccardi considered the work of his pupil a more accurate and substantial scientific contribution to knowledge – with respect to its and to Campedelli's formulation – "that closes and clarifies the concept of *perpendicularity* also in axonometric projection" [12]. So it was not, because the use of polarity rendered the formulations not suitable for those fasting from an in-depth knowledge of projective geometry. In this regard, it should be remembered that in the Italian Faculties of Engineering and Architecture, its teaching was no longer mandatory, and

students do not chose it in the curriculum; soon, it definitively disappeared from the universities teaching's offer. Moreover, the formulated graphic constructions were extremely heavy and complex, also the simplest among them.

So much so that, only two years later, another member of the same school published a new contribution in this regard, the purpose of which was "the determination of the perpendicularity conditions in axonometric projection, with its constructions, without resorting to the polarity" [9]. In fact, is reached quite simple formulations, but the related graphic constructions are not just as, anyway requiring rebatments of planes or switching to the orthographic projections; also in the cavalier projection that, as a particular case of the oblique axonometric projection, it is always easier.

The studies were not particularly development. *Scienza della rappresentazione*, by Mario Docchi and Riccardo Migliari, which is a kind of synthesis of the Descriptive Geometry state at the last century end, does not record anything new about it [5]. And neither in *La rappresentazione geometrica dell'architettura*, published in the same period by Anna Sgross, the question is addressed in a comprehensive manner. Here indeed the solution of metric problems, and among these is the perpendicularity between straight line and plane, is postponed to the Monge method, finally "... to transfer the results in the axonometric image by homological transformations ..." [13].

Later, with the emergence of computerized representation, and the neglect of the Descriptive Geometry studies, this subject has been forgotten; no one has been more dedicated. Even those who in recent years, are committed, in Italy, in the renewal of Descriptive Geometry, typically use the rebatment to solve perpendicularity problems in axonometric projection, oblique as well [4].

A few years ago, basing on the intuition of a student of Salerno University's Civil Engineering, our group has developed an unusual, interesting and very simple formulation of the perpendicularity condition, between straight line

and plane, in isometric cavalier projection, precisely in "military axonometric projection". This condition completely disregards polar and anti-polar, rebatments of planes and switching in Monge method [3]. Its geometrical proof, moreover, is very simple and the identification of the perpendicular to a generic plane requires only three steps.

The formulated perpendicularity condition, moreover, is valid for any type of isometric cavalier projection; or rather when the projection plane coinciding with one of the reference's planes, and the projectors is an oblique direction, inclined by 45° to the projection plane.

2. THE ISOMETRIC "MILITARY" PROJECTION MODEL¹

The geometric model, that represents the isometric military cavalier projection, is particularly simple if we do not consider the objects' precise position in three dimensional space, with respect to the tri-rectangular triad's axes of reference (O, x, y, z). Or if we consider, instead, only the objects morphology and the positions between them. In this case, indeed, through a simple affine homology, is possible to obtain the axonometric image, only starting from the orthogonal projection on the projection plane.

It may seem like a limitation in the model definition, but it is hardly necessary to recall that even Gaspard Monge developed the Descriptive Geometry regardless of the each object's precise position in space, and developed his method as a method of double orthogonal projection (not triple). The two images of an object are, in fact, sufficient to describe the object and infer its properties, in accordance with the discipline's objectives. As the Master said in its *Leçons' Programme*, "Cet art a deux objets principaux. Le premier est de représenter avec exactitude, sur des dessins qui n'ont que deux dimensions, les objets qui en ont trois, et qui sont susceptibles de définition rigoureuse. [...] Le second objet de la Géométrie descriptive est de déduire de la description exacte des corps tout ce qui suit nécessairement de leurs formes et de leurs positions respectives" [8].

The cavalier military axonometric model is defined:

- assuming that the projection plane π coincides with the xy horizontal plane of reference;
- performing on π two projections, both from at infinity points: the first (S'_∞) perpendicular to the projection plane (so $S'_\infty \equiv Z$), the other by a centre S''_∞ (axonometric direction) suitably chosen.

For an isometric representation in the reference axes' directions, the axonometric projector S''_∞ should have an inclination of 45° on π . Each element of the space is represented on the projection plane by two images, respectively coincident with the orthogonal and the axonometric projection; for the generic point X the distance between the two images X' (orthographic projection) and X'' (axonometric projection) gives the point's distance from π (point's elevation).

The two images of a generic plan α , α' (orthogonal projection) and α'' (axonometric projection), since projections of the same plane on another (π) by two distinct projectors, are related in an homologous correspondence, whose axis is the α plan's trace on π ; the centre is, instead, the projection of one projector from the other, on the same projection plane.

In order to have effective representations, as regards the perception and the evocation of the space, it is advisable that the axonometric direction S''_∞ coincides with the bisector's direction of one of the right angles formed by one of the two axes which identify the projection plane with the third axis, namely the bisector of xz o yz .

In the following, it is assumed that S''_∞ is the bisector's direction of the yz angle; so, the z -axis has axonometric projection z'' overlapped to y (with $z''^+ \equiv y'$) and in true size. In this case, the plane containing x and S''_∞ is bisector (let's call it β_1) of the first and third dihedral angle, identified by the projection plane and the xz -plane; this can be taken as the vertical reference plane, it being thus the yz -plane identified as of the profile plane; while the plan containing $x \perp S''_\infty$ is bisector (let's call it β_2) of the second and

fourth dihedral angle, always identified by the projection plane and the xz -plane.

Moreover, since both projectors are points at infinity and $S'_\infty \equiv Z_\infty \perp \pi$, the homology's centre that represents a generic plane is at infinity too: this point is called Z_∞'' and is the point at infinity of the axonometric projection of z -axis direction. So, in isometric military projection, a plane is represented by an affine homology:

$$\omega_\alpha \equiv \{t_\alpha, Z_\infty'', X'_\alpha, X''_\alpha\}.$$

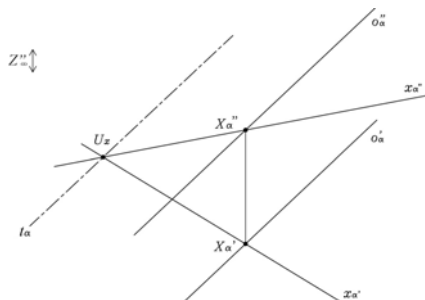


Figure 03. Representation of a generic plane (α) and of its two straight lines (x generic and o horizontal), in the proposed model

The π plane projection is a plane of fixed points, of course isotropic and isometric. Planes parallel to π are represented by a parallel translation and are isotropic and isometrics for both images. Straight lines parallel to π have images parallel to each other and parallel to the traces of planes that contain them.

Vertical planes, containing the first projector, are projecting in the orthogonal projection and are represented by a homology with ratio $k = 0$: $\alpha' \equiv t_\alpha$. Between the vertical planes, those frontal have $t_\alpha \perp Z_\infty''$. Vertical straight lines degenerate in a point in orthogonal projection and have the axonometric projection parallel to z'' .

Planes perpendicular to xz -plane, but not horizontal, being parallel to y -axis, are represented by a singular affine homology ($t_\alpha \perp x$, then contains Z_∞'').

Profile planes are represented by a doubly degenerate homology, with its axis parallel to the Z_∞'' .

In 45° inclined planes on the projection plane, the distance between the plane's trace and the orthogonal projection of an its generic point is equal to the elevation of the considered point (therefore, as the distance between the two images X' and X'').

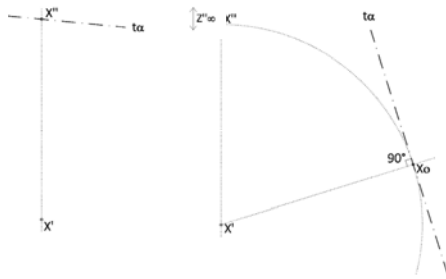


Figure 04. On the left: representation of a plane that contains the axonometric projector S''_∞ ; on the right: a 45° inclined plane on the projection plane

Between the 45° inclined planes on the projection plane, those parallel to β_1 plane (bisector of first and third dihedron) have trace perpendicular to Z_∞ and (containing the projector S''_∞) are represented by a degenerate homology with ratio $k = \infty$. Those parallel to β_2 plane (bisector of the second and fourth dihedron) are perpendicular to the axonometric direction and are represented by an affine orthogonal homology (they also have the trace perpendicular to Z_∞) with ratio of $1/2$, because the orthogonal projection of their generic point is the midpoint of the perpendicular segment from the same point's second image, to the plane's trace.

To increase the effectiveness of solids' representation, with edges parallel to the x and y axes as identified, it is desirable to prevent that the axonometric projection of the axis perpendicular to the projection plane coincides with one of the other axes ($z'' \equiv y'$, with the hypothesis made, or $z'' \equiv x'$, if you choose S''_∞ as the direction of the xz angle's bisector). To this end, can be operated, without any problem, a change of the tri-rectangular triad's axes of reference.

The geometrical model, in fact, solely derives from the choice of the projection plane and of the projectors (S'_∞ and S''_∞), not by the triad of reference from which we can therefore untying. Any pair of lines of π , between them perpendicular, can be taken to the horizontal axes of a tri-rectangular triad's axes of reference, the vertical axis of which is perpendicular to π in their intersection point. Maintaining in O the axes' origin, the vertical axis of the new tri-rectangular triad's axes of reference does not change.

If we assume the same axonometric direction S''_∞ , namely that projects z in the same z'' first identified, the representation of a generic plan – that is the homology which represents it – remains unchanged. In the present case, does not change that of vertical planes, of those projecting into the second projection, of those inclined at 45° on the plane projection; do not change even the representations of frontal planes, of profile planes, of bisecting plans β_1 and β_2 first considered, which however will not be more frontal, of profile and bisector than the new triad.

It follows that the graphic constructions related to the position problems are the same, including those relating to the perpendicular.

So far nothing new, only summary and settlement of what we already knew.

3. THE PERPENDICULARITY CONDITION¹

Precisely a reflection on the bisecting plane β_2 has allowed us to reach the perpendicularity condition of which it was said.

Indeed, if p is the straight line perpendicular to α plane, any plane containing p is perpendicular to α : even the plane φ_p , which projects p in second projection (namely the axonometric projection) and whose trace, therefore, is p'' . This plane, containing the projector S''_∞ , that is the axonometric direction, is therefore perpendicular to the β_2 plane, bisector of the second and fourth dihedron.

Since it φ_p perpendicular both to α and to β_2 , it is also perpendicular to their intersection line: it is $i_\alpha (\equiv \alpha \cap \beta_2)$, which, as perpendicular to φ_p ,

is perpendicular to all the straight lines of this plane. Also to its trace on the projection plane, which is precisely the p'' .

Taking into account that the orthogonal projection on π of a straight line and the trace on π too of the plane to which the line is perpendicular are perpendicular to each other, the orthogonal projection of i_α (it is i_α') is perpendicular to p'' . So, if $p \perp \alpha$ is $p'' \perp i_\alpha'$.

If we start from $p' \perp t_\alpha$ and $p'' \perp i_\alpha'$ and we want to draw in a generic point of p the perpendicular to α , quickly we prove, with a simple traditional process of descriptive geometry, that this line coincides with p .

It can therefore be concluded that, *in isometric military projection, the necessary and sufficient condition so that a straight line is perpendicular to a plane, is that the straight line's orthogonal projection is perpendicular to the*

plane's trace and the straight line's axonometric projection is perpendicular to the orthogonal projection of the intersection's line between this plane and that one perpendicular to the axonometric direction.

Finding i_α' is simple: β_2 has trace (homological axis) perpendicular to the Z_∞'' and passing through the point whose distance, from the orthogonal image of an its generic point X_{β_2} , is equal to its elevation, but is located on the opposite side of the axonometric projection from the first image of this.

Or rather for the symmetric X_{β_2}'' than X_{β_2}' . The i_α' , in this case, will pass to X_α' (orthogonal image of a point that simultaneously lies on α and β_2) and to the intersection point of t_α and t_{β_2} (the straight line is common to the two planes, so its trace belongs to the traces, and homology axes, of both planes).

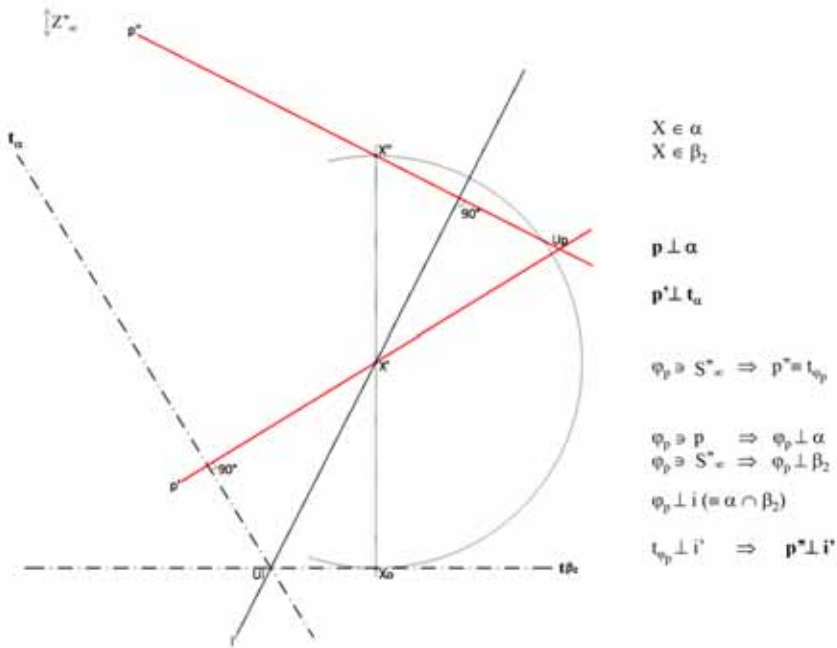


Figure 05. Straight line perpendicular to a generic plane in the proposed model

Note that i_α is precisely the homology's axis that represents the α plan in Monge method, with the same horizontal projection plane, and second reference plane coinciding with the xz -plane (i_α is, in fact, the intersection between α and the bisecting plane of the second and fourth dihedron); its mongean projections will then coincide. It follows that *the second projection, in isometric military projection, of a straight line perpendicular to a plane is perpendicular to the homological axis of the plane in the mongean representation of the same, if the projection plane coincides with that axonometric.*

Based on the necessary and sufficient condition for perpendicularity first formulated – and with appropriate algorithms, as usual grounded on the strictly necessary graphic constructions – it is possible to solve concrete problems. In particular, to draw the p line perpendicular to a given plane, α , in a point X_α of this, we proceed as in the previous figure. Identified the symmetric of X_α'' , with respect to X_α' and in the Z_∞'' direction, from that point we draw the trace (and homological axis) $t_{\beta 2}$; then we join the intersection point, of this and t_α , with X_α' : thus is identified the i_α' . The perpendicular to this line from X_α'' , and the perpendicular to t_α from X_α' are, respectively, the p'' and p' . Instead, to draw a straight line, r perpendicular to a given plane, α , from a point P outside it, it is first convenient to conduct the perpendicular, is still p , to the plane from a point of this – as just seen – and then, from the given point P , the parallel to the so identified line.

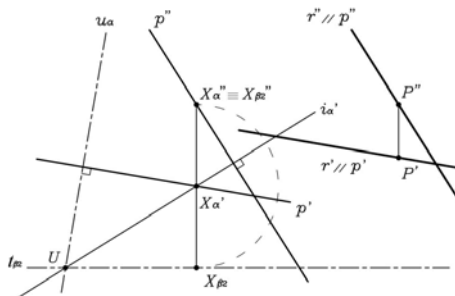


Figure 06. Straight line perpendicular to an assigned plane for an outside point.

To find the plane, α , perpendicular to a straight line r for a given its point, P , we identify the symmetric of P'' , with respect to P' and in the Z_∞'' direction, and from that point we draw the trace (homological axis) $t_{\beta 2}$. Then, conducted by P' the perpendicular to r'' , namely the i_α' , and determined the intersection point of this with $t_{\beta 2}$, t_α is the perpendicular to r' from the so found point.

Instead, to find the plane, β , perpendicular to a line, r , for a given point outside it, Q , it is first convenient to find the trace of the plane perpendicular to r in a generic point of this – as soon as done – and then from the given point Q conduct the plane parallel to that so identified (using, for example, a frontal straight line for Q , of β plane, which is parallel to the frontal straight line of the plane perpendicular to r in its generic point for said point).

This condition, as said, is of general validity for the isometric cavalier projection, not only for the military projection.

In the geometrical model adopted for the isometric military projection, the condition of perpendicularity, between line and plane, is very simple to formulate in the particular case, very recurrent in the representation of architecture and territory, in which the plane is vertical.

The straight lines perpendicular to it are indeed horizontal and, therefore, have the images

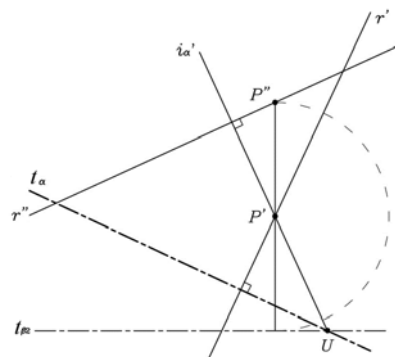


Figure 07. Plane perpendicular to an assigned straight line.

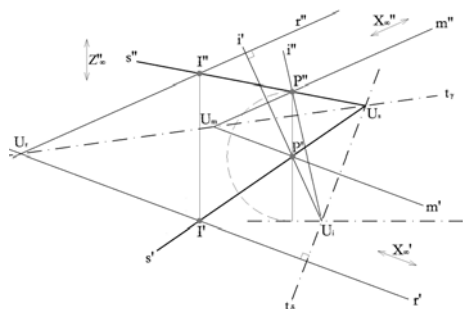


Figure 08. Straight line for a P point and perpendicular to another straight line

parallel to each other; so, having the first projection (orthographic projection) to be perpendicular to the trace (and homological axis) of the plan, it will also be the second projection (namely the axonometric projection).

Vice versa, immediately it shows that if a line has two parallel images (so it is horizontal) and perpendicular to the vertical plane's trace, the line is perpendicular to the same.

And indeed, if we imagine to draw from a generic point, P , of this line, r , the perpendicular, s , to the given plane, this straight line – basing on the direct condition of perpendicularity – will have the orthogonal projection perpendicular to the plane's trace; moreover, since this plane is vertical, s is horizontal: will therefore have the two images parallel to each other.

Therefore, s , having the same name images coinciding with those of r , coincides with that and, therefore, is perpendicular to the given plane.

4. APPLICATIONS ²

Having formulated this condition, therefore, allows to very quickly solve a series of metric problems, as well as position problems – from the simple identification of the perpendicular's foot to a plane, to the determination of distance

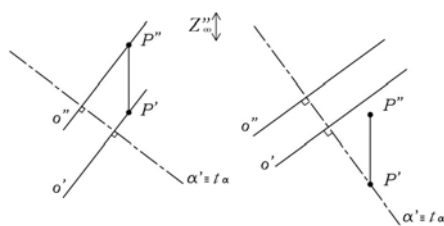


Figure 09. Horizontal straight line, perpendicular to a vertical plane

or angles, to name just a few – and finally simplifies the solids' representation, from the basic to the more complex ones.

More precisely, if we want to represent a right solid with the base belonging to a generic plane, such as a parallelepiped with a rectangular base, the height of which is known, the graphical solution to the problem requires the use of the described perpendicularity's condition.

Drawn, then, the generic α plan whose the base belongs, and identified, by appropriate rebatments, the two projections of the latter, starting from its true size and shape, to complete the solid's representation must draw the p line, perpendicular to α , on which is the solid's height. To do this, find the point of α for which we have to trace the perpendicular (the base's centre or an its vertex, depending on whether you want to represent the axis or an edge of the solid), we proceed using the described methods (cf. § 3).

That is, by matching with the chosen point (in α plane) a point of the β_2 plane, we identify the homological axis of the latter and, consequently, the line of intersection between α and β_2 .

The searched p line will have the first image p' (namely its horizontal projection) orthogonal to the α plane's trace, and the second image p'' (which is properly its axonometric projection) orthogonal to the first image of $i_\alpha (\equiv \alpha \cap \beta_2)$.

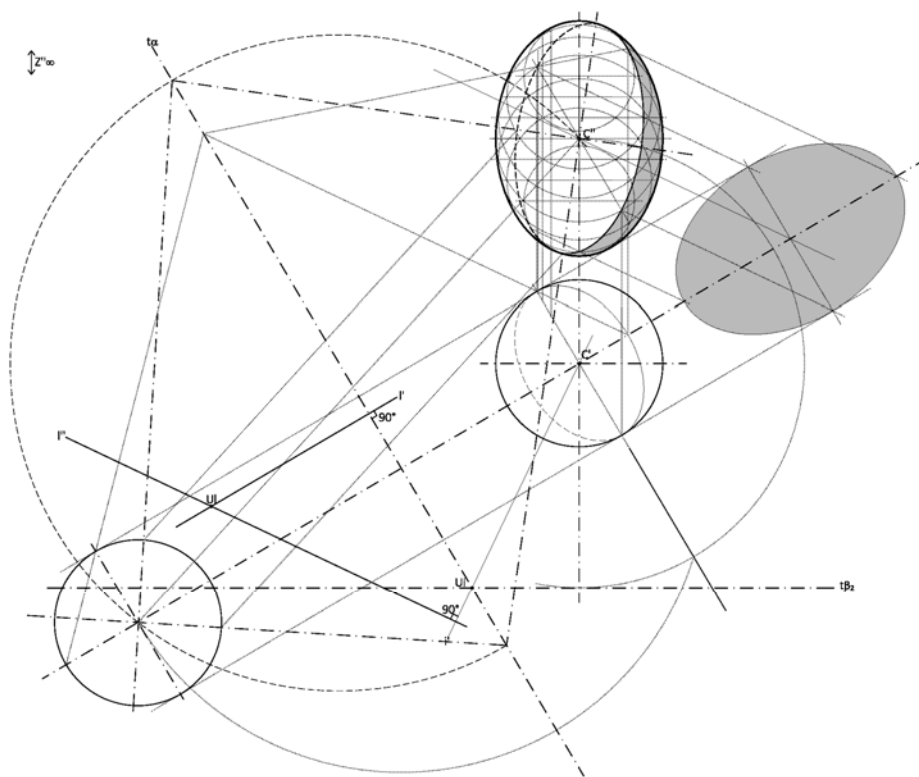


Figure 11. The representation of a sphere's shade

the operational point of view. Indeed, the achieved result does not seem to have any use to the representation made by computer systems, even if it comes to see the possible developments.

Given the traditional difficulty of directly expressing the perpendicularity condition in axonometric projection, it is a very important result, not only because it allows simpler and faster graphical procedures, but mainly because it fills a real emptiness of the isometric model, offered in the courses for architects and engineers, which then appeared unfinished as completely independent model.

REFERENCES

- [1] A. Aterini, *La perpendicolarità in assonometria*, Alinea Editrice, Firenze, 1980.
- [2] L. Campedelli, *Lezioni di geometria. Volume secondo, parte prima: I metodi di rappresentazione della geometria descrittiva*, Cedam, Padova, 1960³.
- [3] V. Cardone, *Modelli grafici dell'architettura e del territorio*, nuova edizione a cura di Salvatore Barba, Cues, Salerno, 2008, pages 128-132 (seconda ri-

stampa: Cues e Maggioli Editore, Sant'Arcangelo di Romagna, 2013).

- [4] L. De Carlo, Rappresentazione in prospettiva parallela (assonometria). In R. Migliari, *Geometria descrittiva*, CittàStudi Edizioni, Novara, 2009, pages 153-205.
- [5] M. Docci and R. Migliari, *Scienza della rappresentazione*, La Nuova Italia Scientifica, Roma, 1992.
- [6] P. Hachette, *Traité de géométrie descriptive: comprenant les applications de cette géométrie aux ombres, à la perspective et à la stéréotomie*, Paris, 1822.
- [7] J.A.R. Maillard de la Gournerie, *Traité de Géométrie Descriptive*, Paris 1860.
- [8] G. Monge, *Géométrie Descriptive*, Baudouin, Paris, an III, 1799.
- [9] A. Pero Nullo, *Alcune note sulla perpendicolarità in assonometria*, Alinea Editrice, Firenze, 1982.
- [10] W.H. Roever, Descriptive geometry and its merits as a collegiate as well as an engineering subject. In *The American mathematical monthly official journal of the mathematical association of America*, vol. XXV, n. 4, april 1918, pages 145-159.
- [11] U. Saccardi, *Applicazioni della geometria descrittiva*, LEF, Firenze, 1977⁴.
- [12] U. Saccardi, *Prefazione a La perpendicolarità in assonometria*, Alinea Editrice, Firenze, 1980.
- [13] A. Sgrosso, *La rappresentazione geometrica dell'architettura*, UTET, Torino, 1996.
- [14] E. Torroja, *Axonometría, o Perspectiva axonométrica*, Madrid, 1879.
- [15] E. Valenzuela, *Axonometría rectangular, o Perspectiva axonométrica rectangular. Perspectiva caballera. Sombras*, Guadalajara, 1896.

ABOUT THE AUTHORS

1. Vito Cardone is full professor of "Drawing", and president of UID (Italian Union of Drawing). His research concerns principally

the descriptive geometry; the elaboration of graphic models of architecture and territory; the history of graphic representation; the history of engineering and of engineer teaching.

2. Barbara Messina is assistant professor of "Drawing". Her research is aimed at the descriptive geometry; the representation as a knowledge system of historical architecture; the drawing in the digital age.

A NEW FOCUS FOR TEACHING AND LEARNING WITH DESCRIPTIVE GEOMETRY'S TUTORIALS THROUGH THE WEB

Germán VALENCIA GARCÍA
Colombia

ABSTRACT: The purpose of this paper is to implement a pedagogical proposal for teaching and learning of the descriptive geometry course through tutorials of three-dimensional animations created with 3D Studio Max software, to be seen in a web site, in an illustrative, detailed, and dynamic manner. A methodology is described to design a master class in a particular subject of descriptive geometry to be seen in an average time of 5 minutes, with constructive and visual process explained step by step. This visual aid can be used as pedagogical support for teachers and students; it can also be used as a bibliographic reference in the curricula that offer this course, and is very useful for e-learning; likewise, serves as a role model for teachers and professionals who want to develop similar projects in various topics and applications.

Keywords: Descriptive Geometry, Education of Descriptive Geometry, Descriptive Geometry Tutorial, e-learning.

1. GENERAL FRAMEWORK

The current teaching of the descriptive geometry course in Colombia, which is offered in most academic programs related to the design, measurement, and construction, is facing an academic decline due to the following reasons:

The continuous curricular reforms that are implemented at universities.

The most common curricular reforms include the implementation of other pre-university courses, such as Introduction to Computer Technology, Fundamentals of Math, Reading academic texts in English, and Reading Comprehension in the first semesters of university study; which has led to the reduction of teaching hours of the descriptive geometry course; it has even reached the point of merging this course with technical drawing to create a new subject in the curriculum.

The lack of spatial reading and blueprints interpretation in secondary education.

Most freshmen have shortcomings of achieved goals in secondary education, due to little or no training in spatial reading and blueprints interpretation, making models, geometric proportions, plane transformations, perspective drawing, scale, and skills in the use of drawing conventional instruments.

Prolonged use of conventional board to explain with extensive drawings.

The explanation of a topic of descriptive geometry using conventional board requires much time and labor to make dihedral projections, perspectives, and notes. The board is a two-dimensional and limited mean, so most students find it difficult to imagine and interpret three-dimensional concepts explained by the teacher in charge.

The speed required to learn and teach with computer advances.

Computer and technological advances that are currently presented require the need to implement audiovisual teaching materials prone to three-dimensional representation, as well as animations and / or virtual simulations; in short, the multimedia technology represents a "lifestyle" that is being imposed in this generation.

The need to implement more educational and illustrative books.

Conventional descriptive geometry books are characterized by their extensive technical and scientific textual explanations of support that a given graph, this graph also shows the approach, the process and result in a single drawing [10], leading to a significant delay in understanding the order in the process of the drawing (approach, constructive process, and outcome); in addition to the inadequate hierarchy of line thickness use and arrowheads, which lead to confusing views and interpretations. Note that most of these books do not update their contents in later editions.

This critical situation is leading to inadequate academic preparation of students for their future professional performance, especially when facing the design of work and solving problems related to the application of descriptive geometry.

The teaching of descriptive geometry by conventional and / or digital resources is still a controversial issue; the most important thing is to experiment with new teaching strategies that facilitate the understanding of this subject during the learning process.

2. PERSONAL REVIEW

The reasons explained above together with the lack of support of educational resources, the reduction of the descriptive geometry curriculum, the inappropriate manipulation of teaching schedules and the null training stimuli during my teaching experience (1997-2006) at the school of Civil engineering and Geomatics of University of Valle, and in the Department of Architecture and Design at the National University of Colombia, forced me to give up the workforce. Being in this discontent I decided to write my book "Descriptive Geometry, step by step" [9], which served me as my motivation and bibliographic reference to begin to develop each topic of it using three-dimensional animations (videos).

Taking advantage of my free time and on my own initiative I decided to self-train myself with books related to the management of software "3D Studio Max" [1-4], [6], [8] and "AutoCAD" [7], with the purpose of teaching and provide a useful intellectual contribution to humanity.

3. METHODOLOGICAL PROPOSAL

The main objective is to implement a visual aid to serve as an instructional support for teaching and learning in each subject or topic of the descriptive geometry course, in an illustrative, detailed and dynamic way, which can be displayed on a website.

With the powerful tools provided by the 3D Studio Max software (modeling, modifiers, mapping, lighting, camera operation, animation and rendering); it is possible to design a video clip of a master class of a certain subject to be seen in an average of 4-5 minutes, considering the following aspects:

3.1 Combining the use of conventional board with the models.

These two educational resources can be fused to conceive the main idea of creating a new dynamic pedagogical strategy that includes all geometric components that take part in a system of dihedral projection, supported by step by step text descriptions, Figure 1.

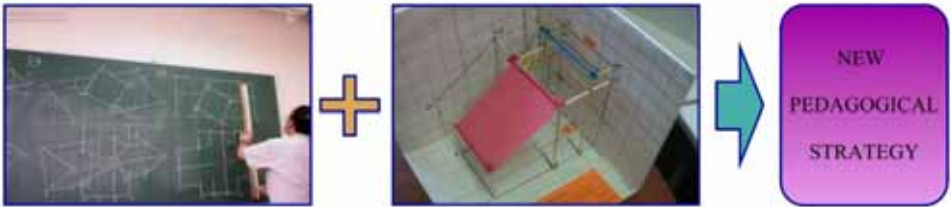
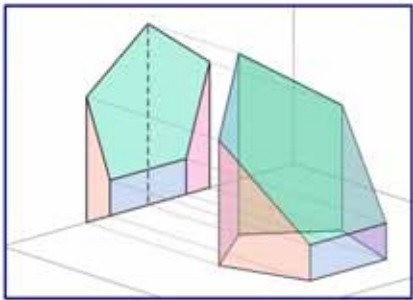


Fig. 1: Using the conventional board and three-dimensional models.

3.2 Show the dynamics of dihedral projection.

The dihedral projection theory should be understood as a system of representation in constant motion, where all the components interact in an orderly manner to show the projection of the object (s) of study, obtaining the views, the relationship between adjacent views, and the folding of the projection planes, Figure 2.



3.3 Demonstrate the theorems according to the case.

If there are involved one or more theorems in a particular theme of descriptive geometry, they must be textually stated , and demonstrated in a simple, illustrative and didactic way and in the right moment in the three-dimensional animation that is being developed, Figure 3.

3.4 Show the construction processes of geometric shapes.

Most geometric objects involved in a three-dimensional scene require a specific constructive process of transformation (Shift, Revolution, Extrude, Sweep, Deformation, Extension, etc.), that should be displayed in a dynamic sequence including its respective textual explanation, Figure 4.

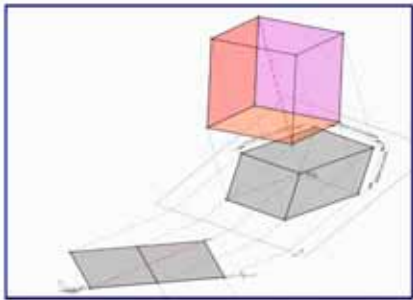


Fig. 2: Dynamic of the dihedral projection (Projection and folding of the projection planes)

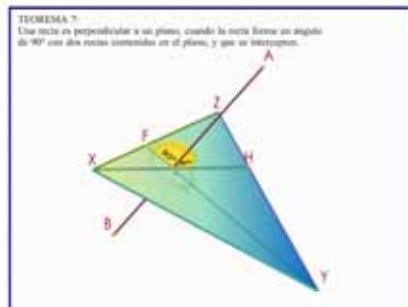
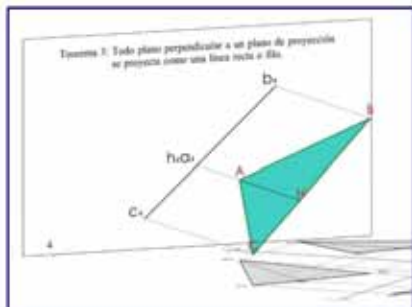


Fig. 3: Demonstration of two cases of theorems

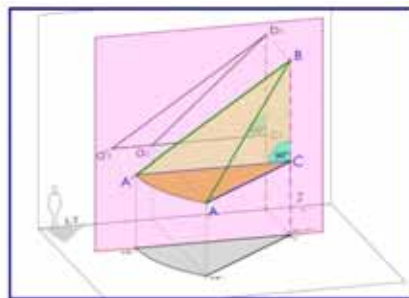
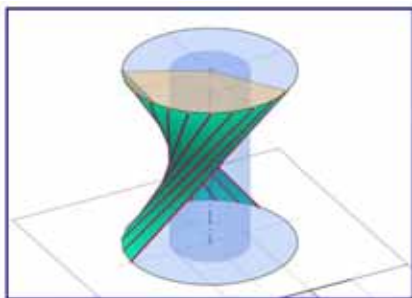


Fig. 4: Description of two construction processes.

3.5 Integrate the viewer to the dihedral projection system of 3D scene.

In a system of dihedral projection are involved the observer, the geometric object, and the projection plane; therefore, the viewer must integrate itself to the dynamics of the dihedral system of the scene 3D as if it was the direct observer. This role allows a more didactic and real interactivity; since the corresponding fields of vision to the main and auxiliary geometric projections of the object(s) are obtained.

It is convenient to use the DIN projection system or first quadrant projection system, because the observer visualizes the object in foreground, and the projection planes with their respective views in the background (behind the object) ¹.

¹ Author's video tutorial:
<http://www.youtube.com/watch?v=IOjEge3d2lo>

3.6 Structure the 3D tutorial as a film script.

The design and planning of 3D tutorial should be structured like the script of a movie "Set up, Confrontation, and Resolution" with its characters interacting in a studio; analogously the tutorial storyboard is "approach, descriptive process and outcome", where the objects of study assume the role of leading actors, and the projection box represents the set. Imagination and creativity are important aspects to tell this script, Figure 5.

The no inclusion of audio in 3D tutorials allows freedom of expression about concepts of the subject of study by teachers of the descriptive geometry course; this way you can control your lecture as it best suits your students.

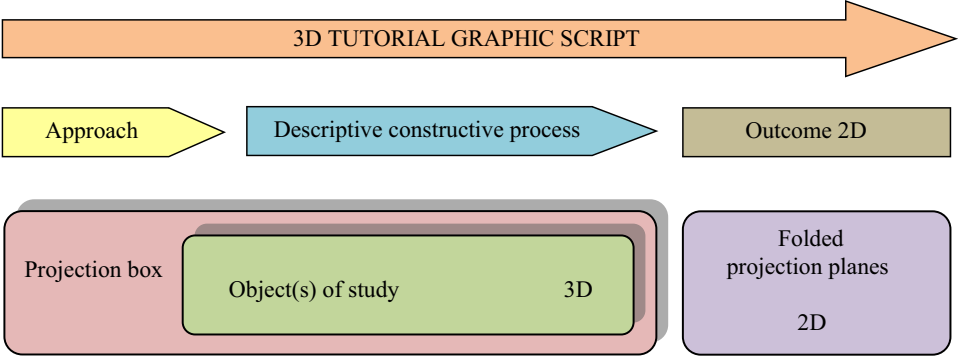


Fig. 5: Structure of the 3D tutorial as a film script.

The implementation of the above aspects is achieved through a combined plan of educational needs focused on precise and clean visual communication, so that allows understanding and relationship of geometric shapes positioned in space.

As the tutorial takes place in a timeline, you must perform an orderly cleaning (Démarche) [5] at appropriate times when required; such cleaning must be consistent and gradual to the criteria of the dihedral projection system, in order to provide the viewer a more didactic and easy understanding of the subject treated.

4. CONCLUSIONS

This didactic audiovisual resource is suitable to be implemented as bibliographic reference in university study plans that offer descriptive geometry course.

The main concept of the subject is conceived as a visual-spatial cognitive impression.

This audiovisual strategy serves as a model to be followed by teachers and professionals who want to develop similar projects in various topics and applications.

Each project can be used as a pedagogical support for teachers and students.

This type of digital resource is ideal for teaching "e-learning".

The 3D animations help facilitate the spatial reading to those people that find it difficult to imagine the spatial perception of geometric forms.

ACKNOWLEDGEMENTS

I thank my wife Claudia Sarmiento and my children Melissa and Nicolas for their constant love and unconditional support, my siblings Jesus and Adriana for their advice; a very special recognition to Mauricio Lenis Sarmiento graduated in Mathematics and Electronics Technologist for his comments and discussions on pedagogy; my sincere gratitude to Jaime Ramirez Pineda graduated in Systems Engineering for his works in Spanish-English translations.

Finally, I thank the developers of the YouTube site for allowing the hosting of my video tutorials of descriptive geometry, and its publishing on the web.

REFERENCES

[1] M, Bousquet. *Trucos con 3Ds Max 2.009*. Primera edición, Alfaomega Grupo Editor, S.A de C.V México, 2.009.

[2] D, Burgos. *3D Studio Max práctico, Guía de aprendizaje*. McGraw-Hill/Interamericana de España, S.A.U, Edificio Valreaty, 1ª planta Basauri, 17 28023 Aravaca, Madrid, 1.997.

[3] J.P, Chismar. *Edición especial 3D Studio Max 3 Animación*. Pearson Educación, S.A. Nuñez de Balboa, 120 28006 Madrid, 2.000.

[4] J.P, Chismar. *Edición especial 3D Studio Max 4 Animación*. Pearson Educación, S.A. Nuñez de Balboa, 120 28006 Madrid, 2.002.

[5] R, Duval. *Semiosis y pensamiento humano*. Peter Lang S.A Editions Scientifiques Europeennes. Page 31. Universidad del Valle, Instituto de Educación y Pedagogía. Grupo de Educación Matemática. Ciudad Universitaria Meléndez, Edificio 388, Oficina 3017. Cali, Colombia. 1.995.

[6] K, Lee. *Edición especial 3Ds Max 4*. Pearson Educación, S.A. Nuñez de Balboa, 120 28006 Madrid, 2.002.

[7] J, López y J.A Tajadura. *Autocad 2.000 Avanzado*. McGraw-Hill/Interamericana de España, S.A.U, Edificio Valreaty, 1ª planta Basauri, 17 28023 Aravaca, Madrid, 1.999

[8] M.T, Peterson. *Descubre 3D Studio Max 3*. Prentice Hall, Iberia, S.R.L. C/Nuñez de Balboa, 120 28006 Madrid, 2.000.

[9] G, Valencia. *Geometría descriptiva, paso a paso*. Primera edición, Bogotá, D.C, editorial Ecoe Ediciones, Colombia, mayo de 2.009. ISBN: 978-958-648-601-9

[10] L, Wellman. *Geometría descriptiva*, pages 112, 114, 126 y 283. Segunda edición, Barcelona, editorial Reverte, S.A, 1.973.

ABOUT THE AUTHOR:

Germán Valencia García is an Architect, graduated from Universidad del Valle, Cali (Colombia). He has worked as a professor in the subjects of Descriptive Geometry, Drawing for Engineering, AutoCAD, Expression Media, Drawing and Construction at five universities: Universidad del Valle, Icesi University, Pontificia Javeriana University, and National University of Colombia Manizales and Palmira sites.

Author of the books "Practical Guide to Engineering Drawing" and "Descriptive Geometry, step by step", expert amateur in the management of software 3D Studio Max and AutoCAD; he is currently developing tutorials about Descriptive Geometry as a hobby for the site YouTube², and his current position is Pastry Micro entrepreneur.

²Author's YouTube site:

www.youtube.com/user/gervallengar

A NEW TRIANGLE THEOREM TO SOLVE THE INVERSE KINEMATICS PROBLEM FOR CHARACTERS WITH HIGHLY ARTICULATED LIMBS

Enrique A. ROSALES¹ and Luis E. FALCÓN-MORALES²

¹Universidad Panamericana, México

²Instituto Tecnológico y de Estudios Superiores de Monterrey, México

ABSTRACT: This paper presents a new geometrical method to solve the inverse kinematics (IK) problem, for characters with highly articulated limbs (HAL). Our method relies on a new triangle theorem, which is the main contribution of this work. This theorem, allows to compute the intersection between two circles, without any sine or cosine computation, and in less time than the classic algebraic approach. We also present the mathematical proof for this theorem, which is valid for any triangle. Applying this new method, we are able to produce complex shapes as spirals or springs, with stable and smooth animations, in redundant kinematic chains (KC).

Keywords: Character animation, geometry, triangle theorem, inverse kinematics, circle-circle intersection, geometry theorem

1. INTRODUCTION

Despite massive amount of computer animation tools, some types of characters still remain very difficult to animate. HAL like an elephant trunk [14] or an octopus arm are very difficult to animate using traditional IK solvers.



Figure 1: A Bézier curve rig for an elephant trunk

To animate this kind of limbs, the most common solution is to use an IK solver based on Bézier curves [2]. The curve has vertices and handles which are moved by the animator, and each link adopts a tangent direction to the curve as seen in Figure (1).

The purpose of this work was to find a new solution for this problem, and our main goal was to produce a tool to easily animate a spiral shape like Figure (2) while keeping curves ratios [11] and smooth movements.

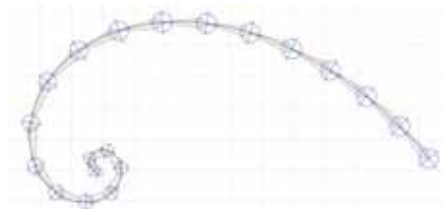


Figure 2: A spiral made by Bézier curves

2. INVERSE KINEMATICS

The IK research in robotics, have its own concerns as inertia, forces, or angular and linear velocities [1, 7, 8]. But this physical concerns can be ignored in animation and computer graphics [10].

So if we work in a virtual world, we only need to calculate the position for each link in the KC. So it is possible to determine our IK problem as equivalent to a *Circle-Circle Intersection Problem* (CCI) [9, 15, 16] for each pair of links, as seen in Figure (3), where the origin denoted as O_1 is the center of the first circle, the end effector

denoted as O_2 is the center of the second circle, the distance between the circles denoted as d , the radii of the circles denoted as r_1 and r_2 representing two links, and P_1 and P_2 as the two possible solutions for the intersection.

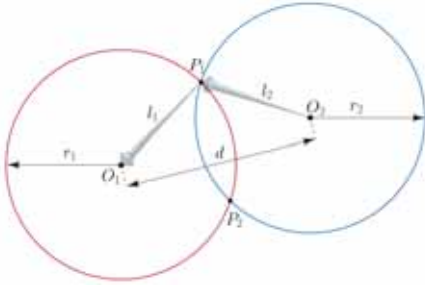


Figure 3: Inverse kinematics as circle-circle intersection

2.1 Singularities

In every KC there are some configurations where the IK problem has no solution, for example: the goal could be unreachable for the end effector, this types of no-solution configurations are known as a singularities.

Given two circles in a two dimension frame of reference, they could have several relative positions and relations of their radii. We defined five possible combinations, as the singularities of our system.

Exterior Singularities

When $d \geq r_1 + r_2$ we have two possible singularities:

- (1) Exterior circles
When $d = \overline{O_1O_2} > r_1 + r_2$, and there are no intersections between the circles.
- (2) Tangent exterior circles
When $d = \overline{O_1O_2} = r_1 + r_2$, and there is only one intersection point between the circles.

In this two cases, the KC must adopt a fully extended posture where the orientation of each

link is the vector $\overline{O_1O_2}$, so no IK compute is needed.

Interior Singularities

When $d \leq |r_1 - r_2|$ we can have three different singularities:

- (3) Interior circles
When $d = \overline{O_1O_2} < |r_1 - r_2|$, and there are no intersections between the circles.
- (4) Tangent interior circles
When $d = \overline{O_1O_2} = |r_1 - r_2|$, and there is only one intersection between the circles.
- (5) Overlapping equal circles
When $O_1 = O_2$ and $r_1 = r_2$ and there are infinite intersections between the circles.

An IK solution is needed for tangent exterior, or tangent interior circles, in this cases, we determine a safe distance that must be larger than $|r_1 - r_2|$, in which the IK solution will be calculated, this distance is a vector with the same direction than $\overline{O_1O_2}$.

3. PROPOSED METHOD

The first step of our method is to find the coordinates of the perpendicular projection from \overline{OJ} towards \overline{OP} , denoted as j as seen in Figure (4), using our new theorem that requires a geometrical construction which is explained as follows:

We draw a circle with a radius equal to l_1 and the center of which is at O , the intersection between this circle and \overline{OP} towards P is denoted by S_1 . Then we draw another circle with a radius equal to l_2 with P as origin, the intersection between this new circle and \overline{PO} towards O is denoted by S_2 .

Now, we determine m as the middle point between S_1 and S_2 .

3.1 The new theorem

Let $\triangle OPJ$ be a triangle like the one in Figure (4), where $A = |\overline{OP}|$, $B = |\overline{Om}|$ and $C = |\overline{S_2S_1}|$, the

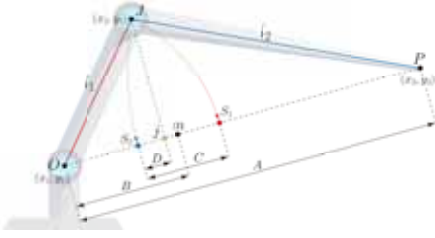


Figure 4: New method construction

distance from S_2 to j (denoted by D) is equal to $\frac{B+C}{A}$, that is:

$$\frac{A}{B} = \frac{C}{D}. \quad (1)$$

The theorem is valid for any configuration of triangles as shown in Figure (5).

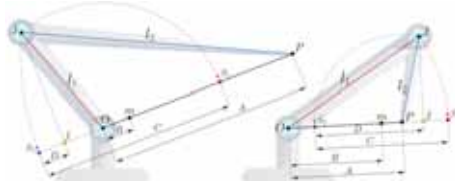


Figure 5: Possible configurations

3.2 Proof

WLOG, we can place the triangle $\triangle OPJ$ on a different direction, matching \overline{OP} with the X positive axis, and O matching the origin of the coordinate system XY . Furthermore, we will use the same notation of the Figure (??) and we will call θ the angle formed by \overline{OP} and \overline{OJ} . If $A \geq l_2$, then S_2 is on the positive side of X axis; however, if $A < l_2$, then S_2 is on the negative side of X axis (Figure 6). To simplify the notation, when we wrote \pm will mean that the sign $+$ is for the case of $A \geq l_2$ and the sign $-$ for the case of $A < l_2$; in the same way, when we wrote \mp will mean that the sign $-$ is for the case of $A < l_2$, and the sign $+$ is for the case of $A \geq l_2$.

$$A = |\overline{OP}| \quad (2)$$

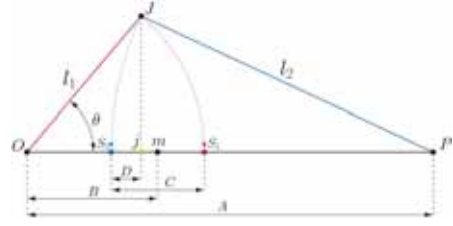


Figure 6: Theorem proof

$$\begin{cases} \overline{OS_1} = l_1 \\ \overline{OS_2} = \pm(A - l_2) \end{cases} \quad (3)$$

$$B = \overline{Om} = \frac{1}{2}[\overline{OS_1} + \overline{OS_2}] = \frac{1}{2}[l_1 + A - l_2] \quad (4)$$

$$C = \overline{S_2S_1} = \overline{OS_1} - \overline{OS_2} = l_1 - (A - l_2) \quad (5)$$

$$D = \overline{S_2j} = l_1 \cos \theta - \overline{OS_2} = l_1 \cos \theta + l_2 - A \quad (6)$$

Therefore:

$$BC = \frac{1}{2}[l_1 + A - l_2][l_1 - (A - l_2)] \quad (7)$$

$$BC = \frac{1}{2}[l_1^2 - (A - l_2)^2] \quad (8)$$

$$BC = \frac{1}{2}[l_1^2 - A^2 + 2Al_2 - l_2^2] \quad (9)$$

$$AD = A[l_1 \cos \theta + l_2 - A] \quad (10)$$

$$AD = Al_1 \cos \theta + Al_2 - A^2 \quad (11)$$

On the other hand, by the law of cosines:

$$l_1^2 + A^2 - 2Al_1 \cos \theta = l_2^2 \quad (12)$$

From which we can infer:

$$l_1^2 + A^2 - l_2^2 = 2Al_1 \cos \theta \quad (13)$$

$$l_1^2 - A^2 + 2Al_2 - l_2^2 = 2Al_1 \cos \theta + 2Al_2 - 2A^2 \quad (14)$$

$$\frac{1}{2}[l_1^2 - A^2 + 2Al_2 - l_2^2] = Al_1 \cos \theta + Al_2 - A^2 \quad (15)$$

Now, by Equation (9) and Equation (11), it follows that $BC = AD$, which is equivalent to Equation (1). Q.E.D.

Obtaining J

Thus, from this theorem $AD = BC$, and because $D = \vec{j} - \vec{S}_2$, $A = |\vec{OP}|$, $B = |\vec{Om}|$ and $C = \vec{S}_1 - \vec{S}_2$, then we can have an expression for the vector \vec{j} as:

$$\vec{j} = \frac{|\vec{Om}|}{|\vec{OP}|}(\vec{S}_1 - \vec{S}_2) + \vec{S}_2. \quad (16)$$

\vec{Oj} is the projection of \vec{OJ} over \vec{OP} , so the grey triangle $\triangle OjJ$ in Figure (7) is a right-angled triangle of which we know the hypotenuse denoted by l_1 and, from Equation (16), the opposite leg to the angle θ denoted by $|\vec{Oj}|$.

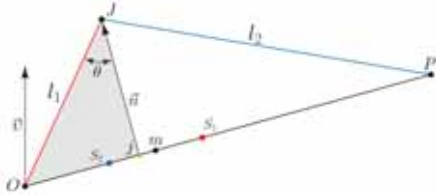


Figure 7: Solving J

With this data, we compute the magnitude of the adjacent leg to the angle θ by the Pythagorean theorem as follows.

$$|\vec{n}| = \sqrt{(l_2)^2 - |\vec{Oj}|^2} \quad (17)$$

The last step to find J is to calculate the vector \vec{n} which must be perpendicular to \vec{OP} with magnitude given by Equation (17) (Figure 7). This

orthogonal vector \vec{n} is obtained with two cross products by the following formula:

$$\vec{n} = (\vec{OP} \times \vec{v}) \times \vec{OP} \quad (18)$$

where \vec{v} is a vector not parallel to OP given by the user and which defines the solving plane. Finally we project the adjacent leg to the angle θ in the direction of \vec{n} , where we can choose whether to use a positive or negative direction to obtain J at the top or bottom, thus:

$$\vec{n} = \sqrt{l_2^2 - |\vec{P} - \vec{j}|^2} \left(\frac{\vec{u}}{|\vec{u}|} \right) \quad (19)$$

And finally, from Equations (16) and (19): $\vec{J} = \vec{j} + \vec{n}$. This method can be implemented in any computer system capable of performing basic vector operations.

4. SOLUTION FOR HAL

In the case of HAL [20] where there are more than two links on the KC, the solution inspired by that of Jamali [3], is to create $n - 2$ virtual links denoted as V_1, V_2, \dots, V_{n-2} given n links l_1, l_2, \dots, l_n , (Figure 8). We use this virtual links to compute the IK problem for the last link denoted as l_n and the first virtual link denoted by V_1 . The solution using Equations (16) and (19), and denoted as J_1 will be the end effector for the next IK computation for V_2 and l_{n-1} . This process is repeated until we have the last two links denoted by l_1, l_2 to obtain the solution J_{n-1} .

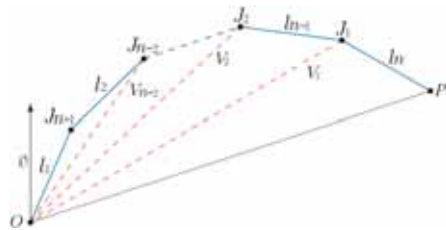


Figure 8: Solution for n links.

In order to use this virtual links denoted by V_k , we need to compute them as prismatic ones with automatic variable length, with the following process: For each link, we define a minimum length denoted by $V_{k_{min}}$ and a maximum length denoted by $V_{k_{max}}$, where the minimum length will be zero and the maximum will be compute with the following equation:

$$V_{k_{max}} = \sum_{i=1}^{n_k} l_i \quad (20)$$

Then, we will need to define a minimum distance from the origin O to the end effector P as \overline{OP}_{min} with any distance > 0 . And a maximum distance from the origin O to the end effector P as \overline{OP}_{max} given by:

$$\overline{OP}_{max} = \overline{V}_{k_{max}} \quad (21)$$

Thus, we can define the length for each virtual link V_k with the following formula:

$$|V_k| = (V_{k_{max}} - V_{k_{min}}) * \left(\frac{(|\overline{OP}| - \overline{OP}_{min}) * 100}{\overline{OP}_{max} - \overline{OP}_{min}} \right) + V_{k_{min}} \quad (22)$$

5. THE SPIRAL SHAPE

In order to obtain an spiral shape like Figure (2), we gradually decrease the length of the virtual vectors by a given factor denoted by δ :

$$|V_k| \leftarrow |V_k| - \delta \quad (23)$$

Where $|V_k| > 0$. This operation, increases the angle of each pair of links, until obtaining the spiral shape as seen in Figure (9).

6. RESULTS

To demonstrate the benefits of the new method, the spiral shaped example was implemented with 50 links. The platform used was Autodesk 3dsMax, using maxscript. This computational analysis was intended to test the transformational in-

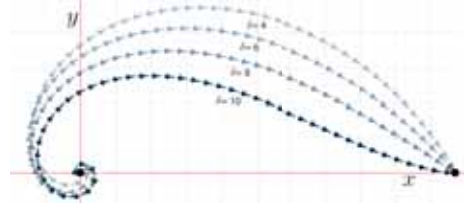


Figure 9: Obtaining the spiral shape

variance, understood as the stability of the solution under the origin and the end effector transformations.

6.1 Native IK solvers.

In Figures (10, 11) we show examples of IK using the native IK solvers from 3dsMax, the IK is resolved, but the way the links are positioned, is far from the aesthetics would be expected for an animation production.

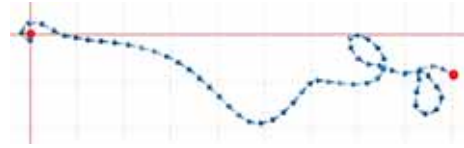


Figure 10: IK solved with History Dependent IK Solver from 3dsMax

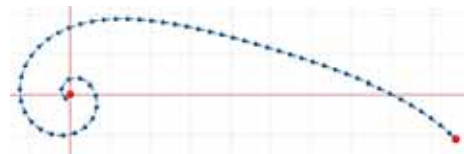


Figure 11: IK solved with History Independent IK Solver from 3dsMax

6.2 The new method result.

According to results, we can infer that the new method is a stable approach to solve the IK problem for HAL, and does not have any sign changes so the IK solution is the same regardless any origin or end effector transformation. Also, the new

method presents a continuous and smooth result (Figure 12).

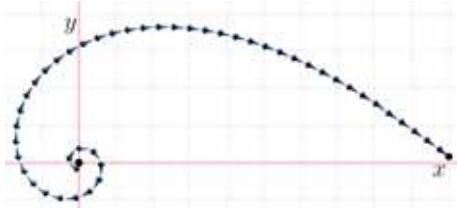


Figure 12: Result with the new method

Also, the method provides an unique solution for each state which is consistent with previous and subsequent states, giving an smooth movement on animation, although the solution for each state is computed independently of the others.

6.3 Results on three dimensions

Using three dimensions coordinates, the new IK solution remains stable during transformations at any quadrant, and the shape and movements are smooth. It is even possible to make a spring shape as seen in Figure (13).

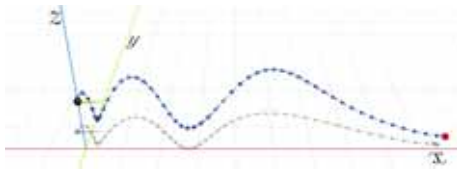


Figure 13: Resulting spring by modifying the \vec{n} vector orientation.

7. CONCLUSIONS

The IK solvers are useful and important tools in animation, it is almost impossible to find a character rig without some sort of IK solver [18], and as the animation industry grows, more complex characters are designed bringing new problems for computer science.

Among the available IK solvers, the spline IK solver is currently the one who gives the best control over curves, while solving the IK problem,

but with great effort from the animator that needs to move the vertices and handles of cubic Bézier curves [11].

Many research has been done about IK solutions for articulated and redundant KCs [3–8, 10, 12, 13, 17, 19, 20], but none of them from the sole animation perspective.

In this paper, we proposed a new method for computing the IK problem for highly articulated limbs. Our method solves an important problem in computer animation, because it allows an animator to create smooth and stable animations of complex curves like the ones of the spider concept shown in Figure (14).

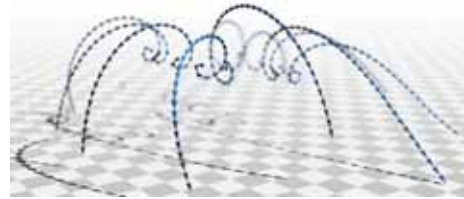


Figure 14: Resulting spring by modifying the \vec{n} vector orientation.

8. CONTRIBUTIONS

The main novelty of this paper, is the new triangles theorem which allows to obtain a circle-circle intersection, with fast and reliable math. The use of this theorem, ensures an stable result under any 3D transformation.

In addition, this work provides an important contribution to the entertainment industry by enabling the making of animations that are currently difficult to obtain with the available animation software .

8.1 Future work

The triangles theorem we proposed, can be used for many other problems such as voronoi diagrams, the pulley problem, or any problem that could need a circle-circle intersection solution.

ACKNOWLEDGMENTS

This work has greatly benefited from discussions with Guillermo A. Parra, Gabriel Castillo, Alejandro García, Joel C. Huegel, Luis E. Mercado, Josué S. Reynoso, Raquel Ruiz, Gildardo Sánchez-Ante, Sergio Velázquez and Julio C. Zamora. To all of them, our thanks and admiration. Also, this work has been partially supported by IJJ, Universidad Panamericana and Instituto Tecnológico y de Estudios Superiores de Monterrey.

REFERENCES

- [1] C. T. Adrià Colomé. Redundant inverse kinematics: Experimental comparative review and two enhancements.
- [2] J. I. T. Alexandre Derouet-Jourdan, Florence Bertails-Descoubes. Stable inverse dynamic curves. *ACM Transactions on Graphics*, 2010.
- [3] M. R. Annisa Jamali, Raisuddin Khan. A new geometrical approach to solve inverse kinematics of hyper redundant robots with variable link length.
- [4] A. D. Aurel Fratu, Laurent Vermeiren. Using the redundant inverse kinematics system for collision avoidance.
- [5] R. F. B. Benhabib, A.A. Goldenberg. A solution to the inverse kinematics of redundant manipulators.
- [6] J. W. Burdick. On the inverse kinematics of redundant manipulators: Characterization of the self-motion manifolds. *Advanced Robotics*, 1989.
- [7] W. K. C. Chevallereau. A new method for the solution of the inverse kinematics of redundant robots.
- [8] C. W. W. I. Daniel R. Baker. Some facts concerning the inverse kinematics of redundant manipulators.
- [9] M. A. J. J. Fatemeh Nourami. Improved circles intersection algorithm for localization in wireless sensor networks.
- [10] T. G. Kang. *Solving Inverse Kinematics Constraint Problems for Highly Articulated Models*. Master's thesis, University of Waterloo, 2000.
- [11] R. L. Levien. *From Spiral to Spline: Optimal Techniques in Interactive Curve Design*. Ph.D. thesis, University of California, Berkeley, 2009.
- [12] C. Q. H. W. Li Sheng, Wang Yiqing. A new geometrical method for the inverse kinematics of the hyper-redundant manipulators.
- [13] A. E. Mehmet, Bodur. Redundant manipulator for obstacle avoidance and inverse kinematics solution by least squares.
- [14] I. D. W. Michael W. Hannan. Kinematics and the implementation of an elephants trunk manipulator and other continuum style robots. *Journal of Robotic Systems*, 2003.
- [15] A. E. Middleditch. Intersection algorithms for lines and circles. *ACM Transactions on Graphics*, 1989.
- [16] C. G. J. P. Robert W. Sumner, Matthias Zwicker. Mesh-based inverse kinematics.
- [17] B. A. J. I. W. Srinivas Neppalli, Matthew A. Csencsits. A geometrical approach to inverse kinematics for continuum manipulators.
- [18] D. C. Tim McLaughlin, Larry Cutler. Character rigging, deformations, and simulations in film and game production.
- [19] C. Welman. *Inverse Kinematics and Geometric Constraint for Articulated Figure Manipulation*. Master's thesis, Simon Fraser University, 1993.

- [20] S. Yahya, H. A. F. Mohamed, M. M. S., and S. Yang. A new geometrical inverse kinematics method for planar hyper redundant manipulators. *Conference on Innovative Technologies in Intelligent Systems and Industrial Applications*, 2009.

ABOUT THE AUTHORS

1. Enrique A. Rosales is Director of the Digital Animation Engineering Program at Universidad Panamericana in Guadalajara, México. erosales@up.edu.mx
2. Luis E. Falcón-Morales is Director of the Computer Science Master Degree program at Instituto Tecnológico y de Estudios Superiores de Monterrey in Guadalajara, México. luis.eduardo.falcon@itesm.mx

NULL POLARITIES AS GENERATORS OF THE PROJECTIVE GROUP

Daniel KLAWITTER

Technical University of Dresden, Institute of Geometry, Germany

ABSTRACT: It is well-known that the group of regular projective transformations of $\mathbb{P}^3(\mathbb{R})$ is isomorphic to the group of projective automorphisms of Klein's quadric $M_2^4 \subset \mathbb{P}^5(\mathbb{R})$. We introduce the Clifford algebra $\mathcal{Cl}_{(3,3)}$ constructed over the quadratic space $\mathbb{R}^{(3,3)}$ and describe how points on Klein's quadric are embedded as null vectors, *i.e.*, grade-1 elements squaring to zero. Furthermore, we discuss how geometric entities from Klein's model can be transferred to this homogeneous Clifford algebra model. Automorphic collineations of Klein's quadric can be described by the action of the so called sandwich operator applied to vectors $v \in \wedge^1 V$. Vectors correspond to null polarities in $\mathbb{P}^3(\mathbb{R})$. We introduce a factorization algorithm. With the help of this algorithm we are able to factorize an arbitrary versor $g \in \mathcal{Cl}_{(3,3)}$ into a set of non-commuting vectors $v_i \in \wedge^1 V$, $i = 1, \dots, k$, $1 \leq k \leq 6$ corresponding to null polarities with $g = v_1 \dots v_k$. Thus, we present a method to factorize every collineation in $\mathbb{P}^5(\mathbb{R})$ that is induced by a projective transformation acting on $\mathbb{P}^3(\mathbb{R})$ into a set of at most six involutoric automorphic collineations of Klein's quadric corresponding to null polarities respectively skew-symmetric 4×4 matrices. Moreover, we give an outlook for Lie's sphere geometry, *i.e.*, the homogeneous Clifford algebra model constructed with the quadratic form corresponding to Lie's quadric $L_1^{n+1} \subset \mathbb{P}^{n+2}(\mathbb{R})$.

Keywords: Clifford algebra, line geometry, Klein's quadric, null polarity, factorization

1. INTRODUCTION

Klein's quadric denoted by $M_2^4 \subset \mathbb{P}^5(\mathbb{R})$ and sub-space intersections of Klein's quadric are well-known, cf [12] or [13]. The group of regular projective transformations of $\mathbb{P}^3(\mathbb{R})$ is isomorphic to the group of automorphic collineations of Klein's quadric. Moreover, Cayley-Klein geometries can be represented by Clifford algebras, where the group of Cayley-Klein isometries is given by the Pin group of the corresponding Clifford algebra, see [6]. Therefore, we recall fundamental basics of Clifford algebras and recall the homogeneous Clifford algebra model corresponding to Klein's quadric. The relationship between projective transformations and elements of the Pin group is given. The full construction can be found in [8].

2. GEOMETRIC ALGEBRA

General introductions to Clifford algebras can be found for example in [4, 5, 10] and [11].

Definition 1. Let V be a real valued vector space of dimension n . Furthermore, let $b : V \mapsto \mathbb{R}$ be a quadratic form on V . The pair (V, b) is called quadratic space.

We denote the Matrix corresponding to b by B_{ij} with $i \leq j, j \leq n$. Therefore $b(x_i, x_j) = B_{ij}$ for some basis vectors x_i and x_j .

Definition 2. The Clifford algebra is defined by the relations

$$x_i x_j + x_j x_i = 2B_{ij}, \quad 1 \leq i, j \leq n. \quad (1)$$

Usually, the algebra is denoted by $\mathcal{Cl}(V, b)$. By Sylvester's law of inertia we can always find a basis $\{e_1, \dots, e_n\}$ of V such that e_i^2 is either 1, -1 or 0.

Definition 3. The number of basis vectors that square to $(1, -1, 0)$ is called signature (p, q, r) . If $r \neq 0$ we call the geometric algebra degenerated. We denote a Clifford algebra by $\mathcal{Cl}_{(p,q,r)}$.

Remark 1. A quadratic real space with signature $(p, q, 0)$ is abbreviated by $\mathbb{R}^{(p,q)}$.

With the new basis $\{e_1, \dots, e_n\}$ the relations (1) become

$$e_i e_j + e_j e_i = 0, \quad i \neq j.$$

In the remainder of this paper we shall abbreviate the product of basis elements with lists

$$e_{12\dots k} := e_1 e_2 \dots e_k, \text{ with } 0 \leq k \leq n.$$

The 2^n monomials

$$e_{i_1} e_{i_2} \dots e_{i_k}, \quad 0 \leq k \leq n$$

form the standard basis of the Clifford algebra. Furthermore, a Clifford algebra is the direct sum $\bigoplus_{i=0}^n \bigwedge^i V$ of all exterior products $\bigwedge^i V$ of any grade $0 \leq i \leq n$ where $e_{k_1} \dots e_{k_i}$, $k_1 < \dots < k_i$ form a basis of $\bigwedge^i V$. Therefore, a Clifford algebra is a graded algebra and its dimension is calculated by

$$\dim \mathcal{C}\ell_{(p,q,r)} = \sum_{i=0}^n \dim \bigwedge^i V = \sum_{i=0}^n \binom{n}{i} = 2^n.$$

Moreover, the Clifford algebra $\mathcal{C}\ell_{(p,q,r)}$ can be decomposed in an even and an odd part

$$\begin{aligned} \mathcal{C}\ell_{(p,q,r)} &= \mathcal{C}\ell_{(p,q,r)}^+ \oplus \mathcal{C}\ell_{(p,q,r)}^- \\ &= \bigoplus_{\substack{i=0 \\ i \text{ even}}}^n \bigwedge^i V \oplus \bigoplus_{\substack{i=0 \\ i \text{ odd}}}^n \bigwedge^i V. \end{aligned}$$

The even part $\mathcal{C}\ell_{(p,q,r)}^+$ is a subalgebra, because the product of two even graded monomials must be even graded since the generators cancel only in pairs. Elements contained in $\bigwedge^i V$ are called *i-vectors* and the \mathbb{R} -linear combination of *i-vectors* $i \leq n$ is called a *multi-vector*. A multi-vector \mathfrak{A} is called homogeneous if $[\mathfrak{A}]_i = \mathfrak{A}$, $i \leq n$ where $[\cdot]_m$ $m \in \mathbb{N}$ denotes the grade- m part of the multi-vector \mathfrak{A} . The product of invertible vectors is called a *versor* and the product of vectors where at least one vector is not invertible is called a *null versor*.

Definition 4. The center of a ring \mathcal{R} is the set of all elements that commute with all other elements

$$\mathcal{C}(\mathcal{R}) := \{c \in \mathcal{R} \mid cx = xc \text{ for all } a \in \mathcal{R}\}.$$

We are interested in the center of a Clifford algebra, see [5, p.95]. The center of a Clifford algebra $\mathcal{C}\ell_{(p,q,r)}$ is

- (1) $\mathcal{C}(\mathcal{C}\ell_{(p,q,r)}) = \{\alpha e_0 + \beta e_{12\dots n} \mid \alpha, \beta \in \mathbb{R}\}$ if n is odd,
- (2) $\mathcal{C}(\mathcal{C}\ell_{(p,q,r)}) = \{\alpha e_0 \mid \alpha \in \mathbb{R}\}$ if n is even.

For the even part the center is

- (3) $\mathcal{C}(\mathcal{C}\ell_{(p,q,r)}^+) = \{\alpha e_0 \mid \alpha \in \mathbb{R}\}$ if n is odd,
- (4) $\mathcal{C}(\mathcal{C}\ell_{(p,q,r)}^+) = \{\alpha e_0 + \beta e_{12\dots n} \mid \alpha, \beta \in \mathbb{R}\}$ if n is even.

2.1 Clifford Algebra Automorphisms

For our purposes two automorphisms that exist on each Clifford algebra are interesting. The *conjugation* is an *anti-involution* denoted by an asterisk, see [14]. Its effect on generators is given by $e_i^* = -e_i$. There is no effect on scalars. Extending the conjugation by using linearity yields

$$(e_{i_1} e_{i_2} \dots e_{i_k})^* = (-1)^k e_{i_k} \dots e_{i_2} e_{i_1}$$

with $0 \leq i_1 < i_2 < \dots < i_k \leq n$. The geometric product of a vector $\mathbf{v} = \sum_{i=1}^n x_i e_i \in \bigwedge^1 V$ with its conjugate results in

$$\begin{aligned} \mathbf{v}\mathbf{v}^* &= -x_1^2 - x_2^2 - \dots - x_p^2 + x_{p+1}^2 + \dots + x_{p+q}^2 \\ &= -b(\mathbf{v}, \mathbf{v}), \end{aligned}$$

where $\mathbf{v} = (x_1, \dots, x_n)^T \in \mathbb{R}^n$.

Definition 5. The inverse element of a versor $\mathbf{v} \in \mathcal{C}\ell_{(p,q,r)}$ is defined by

$$\mathbf{v}^{-1} := \frac{\mathbf{v}^*}{N(\mathbf{v})},$$

with $N(\mathbf{v}) := \mathbf{v}\mathbf{v}^*$.

The map $N : \mathcal{C}\ell_{(p,q,r)} \rightarrow \mathcal{C}\ell_{(p,q,r)}$ is called the *norm* of the Clifford algebra. For general multi-vectors $\mathfrak{M} \in \mathcal{C}\ell_{(p,q,r)}$ inverse elements exist and are defined through the relation $\mathfrak{M}\mathfrak{M}^{-1} = \mathfrak{M}^{-1}\mathfrak{M} = 1$, but the determination is more difficult and can be found in [3]. Note that in general not every element is invertible. The other automorphism we are dealing with is the *main involution*. It is denoted by α and defined by

$$\alpha(e_{i_1}e_{i_2}\dots e_{i_k}) = (-1)^k e_{i_1}e_{i_2}\dots e_{i_k}$$

for $0 \leq i_1 < i_2 < \dots < i_k \leq n$. The main involution has no effect on the even subalgebra and it commutes with the conjugation, i.e., $\alpha(\mathfrak{M}^*) = \alpha(\mathfrak{M})^*$ for arbitrary $\mathfrak{M} \in \mathcal{C}\ell_{(p,q,r)}$.

2.2 Clifford Algebra Products

On vectors $\mathfrak{a}, \mathfrak{b} \in \bigwedge^1 V$ we can write the inner product in terms of the geometric product

$$\mathfrak{a} \cdot \mathfrak{b} := \frac{1}{2}(\mathfrak{a}\mathfrak{b} + \mathfrak{b}\mathfrak{a}). \quad (2)$$

A generalization of the inner product to homogeneous multi-vectors can be found in [7]. For $\mathfrak{A} \in \bigwedge^k V$, $\mathfrak{B} \in \bigwedge^l V$ the generalized inner product is defined by

$$\mathfrak{A} \cdot \mathfrak{B} := [\mathfrak{A}\mathfrak{B}]_{|k-l|}.$$

There is another product on vectors, i.e., the outer (or exterior) product

$$\mathfrak{a} \wedge \mathfrak{b} := \frac{1}{2}(\mathfrak{a}\mathfrak{b} - \mathfrak{b}\mathfrak{a}). \quad (3)$$

This product can also be generalized to homogeneous multi-vectors, see again [7]. For $\mathfrak{A} \in \bigwedge^k V$, $\mathfrak{B} \in \bigwedge^l V$ the generalized outer product is defined by

$$\mathfrak{A} \wedge \mathfrak{B} := [\mathfrak{A}\mathfrak{B}]_{|k+l|}.$$

From equation (2) and (3) it follows that for vectors the geometric product can be written as the sum of the inner and the outer product

$$\mathfrak{a}\mathfrak{b} = \mathfrak{a} \cdot \mathfrak{b} + \mathfrak{a} \wedge \mathfrak{b}.$$

More general this can be defined for multivectors with the commutator and the anti-commutator product, see [10]. For treating geometric entities within this algebra context the definition of a *k-blade*, the *inner product null space*, and its dual the *outer product null space* is needed.

Definition 6. A *k-blade* is the *k-fold exterior product of vectors* $\mathfrak{v} \in \bigwedge^1 V$. Therefore, a *k-blade* can be written as

$$\mathfrak{A} = \mathfrak{a}_1 \wedge \mathfrak{a}_2 \wedge \dots \wedge \mathfrak{a}_k.$$

If the *k-blade* squares to zero, it is called a *null k-blade*. The inner product null space (IPNS) of a blade $\mathfrak{A} \in \bigwedge^k V$, cf. [10], is defined by

$$\text{NI}(\mathfrak{A}) := \left\{ \mathfrak{v} \in \bigwedge^1 V \mid \mathfrak{v} \cdot \mathfrak{A} = 0 \right\}.$$

Moreover, the outer product null space (OPNS) of a blade $\mathfrak{A} \in \bigwedge^k V$ is defined by

$$\text{NO}(\mathfrak{A}) := \left\{ \mathfrak{v} \in \bigwedge^1 V \mid \mathfrak{v} \wedge \mathfrak{A} = 0 \right\}.$$

Remark 2. The same set can be described with inner product or outer product null spaces. In the non-degenerate case the change between both representations is achieved with the pseudoscalar $\mathfrak{J} = e_1 \dots e_n$

$$\text{NO}(\mathfrak{A}) = \text{NI}(\mathfrak{A}\mathfrak{J}).$$

Later on we work in a homogeneous Clifford algebra model, and therefore, it does not matter from which side we multiply with the pseudoscalar, since the inner product respectively outer product null space is not affected. Moreover, multiplication with the pseudoscalar corresponds to the application of the polarity corresponding to the measure quadric. In the degenerate case the *Poincaré-identity* has to be used, compare to [6].

2.3 Pin and Spin Groups

With respect to the geometric product the units, i.e., the invertible elements of a Clifford algebra denoted by $\mathcal{C}\ell_{(p,q,r)}^\times$ form a group.

Definition 7. The Clifford group is defined by

$$\Gamma(\mathcal{C}\ell_{(p,q,r)}) := \left\{ \mathfrak{g} \in \mathcal{C}\ell_{(p,q,r)}^\times \mid \alpha(\mathfrak{g})\mathfrak{v}\mathfrak{g}^{-1} \in \bigwedge^1 V \text{ for all } \mathfrak{v} \in \bigwedge^1 V \right\}.$$

A proof that $\Gamma(\mathcal{C}\ell_{(p,q,r)})$ is indeed a group with respect to the geometric product can be found in [4]. We define two important subgroups of the Clifford group.

Definition 8. The Pin group is the subgroup of the Clifford group with $N(\mathfrak{g}) = \pm 1$.

$$\text{Pin}_{(p,q,r)} := \left\{ \mathfrak{g} \in \mathcal{C}\ell_{(p,q,r)} \mid \mathfrak{g}\mathfrak{g}^* = \pm 1 \text{ and } \alpha(\mathfrak{g})\mathfrak{v}\mathfrak{g}^* \in \bigwedge^1 V \text{ for all } \mathfrak{v} \in \bigwedge^1 V \right\}.$$

Furthermore, we define the Spin group by $\text{Pin}_{(p,q,r)} \cap \mathcal{C}\ell_{(p,q,r)}^+$

$$\text{Spin}_{(p,q,r)} := \left\{ \mathfrak{g} \in \mathcal{C}\ell_{(p,q,r)}^+ \mid \mathfrak{g}\mathfrak{g}^* = \pm 1 \text{ and } \alpha(\mathfrak{g})\mathfrak{v}\mathfrak{g}^* \in \bigwedge^1 V \text{ for all } \mathfrak{v} \in \bigwedge^1 V \right\}.$$

Remark 3. For non-degenerated Clifford algebras the Pin group is a double cover of the orthogonal group of the quadratic space (V, b) . Moreover, the Spin group is a double cover of the special orthogonal group of (V, b) , see [4].

3. THE HOMOGENEOUS CLIFFORD ALGEBRA MODEL CORRESPONDING TO KLEIN'S QUADRIC

To construct a homogeneous Clifford algebra model we use a vector space as model for the projective space which is the base space of the model. Thus, we use \mathbb{R}^6 as model space for $\mathbb{P}^5(\mathbb{R})$ together with the quadric form b given by the corresponding symmetric matrix

$$B = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}$$

where O denotes the 3×3 zero matrix and I the 3×3 identity matrix. The $\bigwedge^1 V$ subspace is identified with the vector space, and therefore, with

the projective space. The square of a general vector $\mathfrak{v} = \sum_{i=1}^6 x_i e_i$ results in

$$\mathfrak{v}\mathfrak{v} = x_1 x_4 + x_2 x_5 + x_3 x_6.$$

This is the equation of Klein's quadric and a vector corresponds to a point on Klein's quadric if it is a null vector, i.e., a vector squaring to zero. Since the Clifford algebra allows the description of subspaces with the use of blades of higher grade, we can express linear line manifolds within the homogeneous Clifford algebra model, see [8, 9].

3.1 Linear Line Manifolds

Clifford algebras carry the subspace structure of Grassmann algebras together with the metric properties derived with the corresponding Cayley-Klein geometry. Therefore, we construct blades of grade k , $k \leq 5$ and examine the geometric inner product and geometric outer product null spaces.

Two-blades. The exterior product of two vectors corresponds to a two-blade that describes a line P_1^1 in $\mathbb{P}^5(\mathbb{R})$. We use the outer product null space to determine the set of points on P_1^1 contained in M_2^4 , and therefore, the set of lines in $\mathbb{P}^3(\mathbb{R})$ corresponding to the two-blade. If the whole line P_1^1 is contained in M_2^4 the corresponding two-blade is a null two-blade. Thus, the outer product null space corresponds to a pencil of lines in $\mathbb{P}^3(\mathbb{R})$ or to two lines in $\mathbb{P}^3(\mathbb{R})$ that may be identical. The inner product null space of a two-blade corresponds to a three-space intersection with Klein's quadric, i.e., a linear line congruence in $\mathbb{P}^3(\mathbb{R})$.

Three-blades. Three-blades can be generated as the exterior product of three vectors and describe two-spaces in $\mathbb{P}^5(\mathbb{R})$. The outer product null space of a three-blade corresponds to the two-space intersection of the two-space P_1^2 described by the three-blade with M_2^4 . If the three-blade is a null three-blade P_1^2 is completely contained in Klein's quadric and the outer product null space corresponds to a bundle or a field of lines in

$\mathbb{P}^3(\mathbb{R})$. The outer product null space of a non-null three-blade corresponds to a conic section of M_2^4 that defines a regulus in $\mathbb{P}^3(\mathbb{R})$. In a dual way the inner product null space of a three-blade is the intersection of the image of P_1^2 under the polarity corresponding to the measure quadric with Klein's quadric.

Four-blades. Outer product null spaces of four-blades correspond to three-space intersections with Klein's quadric and the corresponding set of lines in $\mathbb{P}^3(\mathbb{R})$ is a linear congruence of lines. The dual of a three-space in $\mathbb{P}^5(\mathbb{R})$ is a line. Thus, the inner product null space of a four-blade is the intersection of this line with M_2^4 .

Five-blades. Five-blades describe four-spaces in Klein's model and the intersection of a four-space with Klein's quadric results in a linear congruence of lines. The outer product null space of a five-blade corresponds to a linear congruence of lines. Moreover, the inner product null space of a five-blade yields a vector.

3.2 Clifford Group Action

The null vectors correspond to points on Klein's quadric. Moreover, the sandwich action of an arbitrary versor on a null vector results a null vector again. Thus Klein's quadric is mapped to itself. To examine the action of the whole Clifford group, we first investigate the action of non-null vectors on null vectors.

Remark 4. *Since we are working in a homogeneous Clifford algebra model multiplication with a real factor does not change the geometric meaning of an algebra element. Therefore, we use the conjugate element instead of the inverse element. Hence, the sandwich operator that we use is given by:*

$$\alpha(a)va^* = aa^*(\alpha(a)va^{-1}).$$

This operator does not involve an inverse, and therefore, it can also be applied if the element a is not invertible.

Sandwich action of vectors. Let $a = \sum_{i=1}^6 a_i e_i$ and $v = \sum_{i=1}^6 x_i e_i$ be two vectors with $aa \neq 0$ and $vv = 0$. The action of the sandwich operator $\alpha(a)va^*$ on the $\wedge^1 V$ subspace can be expressed as product of a matrix with a vector. The matrix has the form:

$$M = \begin{pmatrix} k_1 & a_1 a_5 & a_1 a_6 & a_1 a_1 & a_1 a_2 & a_1 a_3 \\ a_2 a_4 & k_2 & a_2 a_6 & a_2 a_1 & a_2 a_2 & a_2 a_3 \\ a_3 a_4 & a_3 a_5 & k_3 & a_3 a_1 & a_3 a_2 & a_3 a_3 \\ a_4 a_4 & a_4 a_5 & a_4 a_6 & k_4 & a_4 a_2 & a_4 a_3 \\ a_5 a_4 & a_5 a_5 & a_5 a_6 & a_5 a_1 & k_5 & a_5 a_3 \\ a_6 a_4 & a_6 a_5 & a_6 a_6 & a_6 a_1 & a_6 a_2 & k_6 \end{pmatrix}, \quad (4)$$

with

$$\begin{aligned} k_1 &= -a_5 a_2 - a_6 a_3, & k_2 &= -a_6 a_3 - a_4 a_1, \\ k_3 &= -a_4 a_1 - a_5 a_2, & k_4 &= -a_5 a_2 - a_6 a_3, \\ k_5 &= -a_6 a_3 - a_4 a_1, & k_6 &= -a_4 a_1 - a_5 a_2. \end{aligned}$$

This involutoric automorphism of Klein's quadric corresponds to a null polarity acting on $\mathbb{P}^3(\mathbb{R})$, see [8]. Furthermore, the Clifford group is generated by invertible vectors, and therefore, all elements of the Clifford group correspond to compositions of null polarities in $\mathbb{P}^3(\mathbb{R})$. Moreover, this means that every element of $\mathcal{C}\ell_{(3,3)}^+$ that is an even product of vectors corresponds to a collineation and every element of $\mathcal{C}\ell_{(3,3)}^-$ that is an odd product of vectors corresponds to a correlation.

Collineations. A general element $g \in \mathcal{C}\ell_{(3,3)}^+$ corresponding to a collineation is given by

$$\begin{aligned} g &= g_1 e_0 + g_2 e_{12} + g_3 e_{13} + g_4 e_{14} + g_5 e_{15} + g_6 e_{16} \\ &+ g_7 e_{23} + g_8 e_{24} + g_9 e_{25} + g_{10} e_{26} + g_{11} e_{34} \\ &+ g_{12} e_{35} + g_{13} e_{36} + g_{14} e_{45} + g_{15} e_{46} + g_{16} e_{56} \\ &+ g_{17} e_{1234} + g_{18} e_{1235} + g_{19} e_{1236} + g_{20} e_{1245} \\ &+ g_{21} e_{1246} + g_{22} e_{1256} + g_{23} e_{1345} + g_{24} e_{1346} \\ &+ g_{25} e_{1356} + g_{26} e_{1456} + g_{27} e_{2345} + g_{28} e_{2346} \\ &+ g_{29} e_{2356} + g_{30} e_{2456} + g_{31} e_{3456} + g_{32} e_{123456}. \end{aligned}$$

We derive constraints to this element by $\alpha(g)vg^* \in \wedge^1 V$ for all $v \in \wedge^1 V$. If we try to

compute the corresponding collineation we have to distinguish the type of the collineation, *i.e.*, if it acts on points or on planes.

Action on Points. First we assume $\mathfrak{g} \in \mathcal{C}\ell_{(3,3)}^+$ corresponds to a collineation acting on points. The automorphic collineation of Klein's quadric induced by \mathfrak{g} can be transferred to a collineation that acts on $\mathbb{P}^3(\mathbb{R})$. Therefore, the element $g \in \mathcal{C}\ell_{(3,3)}^+$ is applied to a null three-blade corresponding to a bundle of lines under the constraint that \mathfrak{g} is a versor, see [8]. This results in the matrix $(m_{i,j}), i, j = 0, \dots, 3$ for the collineation with

$$\begin{aligned} m_{00} &= g_1 - g_{20} - g_{24} - g_{32} - g_{29} + g_9 + g_4 + g_{13}, \\ m_{11} &= g_{24} - g_9 + g_{20} - g_{13} - g_{32} + g_1 + g_4 - g_{29}, \\ m_{22} &= g_1 - g_{13} - g_{32} - g_4 + g_{29} + g_9 - g_{24} + g_{20}, \\ m_{33} &= g_{24} + g_{13} + g_{29} + g_1 - g_4 - g_9 - g_{20} - g_{32}, \end{aligned}$$

and

$$\begin{aligned} m_{01} &= 2(g_7 + g_{17}), & m_{02} &= 2(g_{18} - g_3), \\ m_{03} &= 2(g_{19} + g_2), & m_{10} &= -2(g_{26} + g_{16}), \\ m_{12} &= 2(g_5 + g_{25}), & m_{13} &= 2(g_6 - g_{22}), \\ m_{20} &= 2(g_{15} - g_{30}), & m_{21} &= 2(g_8 + g_{28}), \\ m_{23} &= 2(g_{21} + 2g_{10}), & m_{30} &= -2(g_{31} + g_{14}), \\ m_{31} &= 2(g_{11} - g_{27}), & m_{32} &= 2(g_{23} + g_{12}). \end{aligned}$$

The system of equations

$$(m_{i,j}) = (c_{i,j}), \quad i, j = 0, \dots, 3 \quad (5)$$

can be used to compute an element of $\mathcal{C}\ell_{(3,3)}^+$ corresponding to a collineation acting on points with matrix representation $C = (c_{i,j}), i, j = 0, \dots, 3$. Therefore, constraint equations derived from $\alpha(\mathfrak{g})\mathfrak{v}\mathfrak{g}^* \in \wedge^1 V$ and $\mathfrak{g}\mathfrak{g}^* = \pm 1$ are used.

Action on Planes. Moreover, the same automorphic collineation of Klein's quadric may correspond to a collineation in $\mathbb{P}^3(\mathbb{R})$ that maps planes to planes. This action can also be represented

as matrix vector product. Therefore, the general element $\mathfrak{g} \in \mathcal{C}\ell_{(3,3)}^+$ is applied to a null three-blade corresponding to a field of lines. The action can be expressed as collineation with matrix $(m_{i,j}), i, j = 0, \dots, 3$

$$\begin{aligned} m_{00} &= g_{32} - g_{20} - g_{13} - g_{29} - g_9 - g_{24} + g_1 - g_4, \\ m_{11} &= g_1 + g_9 + g_{20} + g_{24} + g_{13} - g_{29} + g_{32} - g_4, \\ m_{22} &= g_{20} + g_{29} + g_4 + g_{13} - g_9 - g_{24} + g_1 + g_{32}, \\ m_{33} &= g_9 + g_{24} + g_{29} - g_{13} + g_1 + g_4 - g_{20} + g_{32}, \end{aligned}$$

and

$$\begin{aligned} m_{01} &= 2(g_{16} - g_{26}), & m_{02} &= -2(g_{15} + g_{30}), \\ m_{03} &= 2(g_{14} - g_{31}), & m_{10} &= 2(g_{17} - g_7), \\ m_{12} &= 2(g_{28} - g_8), & m_{13} &= -2(g_{27} + g_{11}), \\ m_{20} &= 2(g_3 + g_{18}), & m_{21} &= 2(g_{25} - g_5), \\ m_{23} &= 2(g_{23} - g_{12}), & m_{30} &= 2(g_{19} - g_2), \\ m_{31} &= -2(g_{22} + g_6), & m_{32} &= 2(g_{21} - g_{10}). \end{aligned}$$

Depending on the action of the collineation, *i.e.*, if it acts on points or planes the elements of the Spin group can be transferred to their matrix representations and vice versa.

Correlations. General elements of the odd part $\mathcal{C}\ell_{(3,3)}^-$ have the form

$$\begin{aligned} \mathfrak{h} &= h_1 e_1 + h_2 e_2 + h_3 e_3 + h_4 e_4 + h_5 e_5 + h_6 e_6 \\ &+ h_7 e_{123} + h_8 e_{124} + h_9 e_{125} + h_{10} e_{126} + h_{11} e_{134} \\ &+ h_{12} e_{135} + h_{13} e_{136} + h_{14} e_{145} + h_{15} e_{146} + h_{16} e_{156} \\ &+ h_{17} e_{234} + h_{18} e_{235} + h_{19} e_{236} + h_{20} e_{245} + h_{21} e_{246} \\ &+ h_{22} e_{256} + h_{23} e_{345} + h_{24} e_{346} + h_{25} e_{356} + h_{26} e_{456} \\ &+ h_{27} e_{12345} + h_{28} e_{12346} + h_{29} e_{12356} + h_{30} e_{12456} \\ &+ h_{31} e_{13456} + h_{32} e_{23456}. \end{aligned}$$

Their action on points or planes of the three-dimensional projective space $\mathbb{P}^3(\mathbb{R})$ can be described by the sandwich action on null three-blades corresponding to two-spaces contained entirely in M_2^4 that correspond to bundles of lines or fields of lines.

Action on Planes. The action of a general element $\mathfrak{h} \in \mathcal{C}\ell_{(3,3)}^-$ applied to a null three-blade corresponding to a two-space contained in Klein's quadric that is the image of a field of lines under the Klein map can be expressed as 4×4 matrix $(m_{i,j}), i, j = 0, \dots, 3$ that determines a correlation with $m_{00} = -2h_7$, $m_{11} = -2h_{16}$, $m_{22} = 2h_{21}$, $m_{33} = -2h_{23}$, and

$$\begin{aligned} m_{01} &= h_9 + h_{13} - h_{29} + h_1, \\ m_{02} &= h_2 - h_8 + h_{28} + h_{19}, \\ m_{03} &= h_3 - h_{27} - h_{18} - h_{11}, \\ m_{10} &= h_{29} + h_{13} + h_9 - h_1, \\ m_{12} &= h_6 - h_{22} + h_{15} + h_{30}, \\ m_{13} &= h_{31} - h_{25} - h_5 - h_{14}, \\ m_{20} &= h_{19} - h_8 - h_2 - h_{28}, \\ m_{21} &= h_{15} - h_{30} - h_6 - h_{22}, \\ m_{23} &= h_4 - h_{20} + h_{32} + h_{24}, \\ m_{30} &= h_{27} - h_{11} - h_{18} - h_3, \\ m_{31} &= h_5 - h_{31} - h_{25} - h_{14}, \\ m_{32} &= h_{24} - h_4 - h_{32} - h_{20}. \end{aligned}$$

Action on Points. The same element $\mathfrak{h} \in \mathcal{C}\ell_{(3,3)}^-$ can be interpreted as correlation in $\mathbb{P}^3(\mathbb{R})$ that acts on points and maps them to planes. This action can be described by a product of a matrix $m_{(i,j)}, i, j = 0, \dots, 3$ with a vector. This matrix can be expressed with the coefficients of $\mathfrak{h} \in \mathcal{C}\ell_{(3,3)}^-$ and has the entries $m_{00} = 2h_{26}$, $m_{11} = 2h_{17}$, $m_{22} = -2h_{12}$, $m_{33} = 2h_{10}$, and

$$\begin{aligned} m_{01} &= h_{32} - h_4 - h_{20} - h_{24}, \\ m_{02} &= h_{14} - h_{31} - h_{25} - h_5, \\ m_{03} &= h_{30} + h_{15} + h_{22} - h_6, \\ m_{10} &= h_4 - h_{32} - h_{24} - h_{20}, \\ m_{12} &= h_{18} - h_{27} - h_3 - h_{11}, \\ m_{13} &= h_2 + h_8 + h_{19} - h_{28}, \\ m_{20} &= h_{31} + h_{14} - h_{25} + h_5, \\ m_{21} &= h_3 - h_{11} + h_{27} + h_{18}, \\ m_{23} &= h_9 - h_{13} - h_1 - h_{29}, \\ m_{30} &= h_{15} - h_{30} + h_6 + h_{22}, \end{aligned}$$

$$m_{31} = h_8 - h_2 + h_{28} + h_{19},$$

$$m_{32} = h_1 - h_{13} + h_{29} + h_9.$$

Depending on the action of the correlation, *i.e.*, if it acts on points or planes the elements of the Pin group can be transferred to their matrix representations and vice versa.

4. A FACTORIZATION ALGORITHM AND ITS APPLICATION

At this point we recall a theorem from [8].

Theorem 1. *Every regular collineation or correlation can be expressed as the product of six null polarities at the most.*

For a proof we refer to [8] or [9]. With the help of a modified version of a factorization algorithm introduced in [10, p. 107] we factorize arbitrary regular 4×4 matrices into the product of six skew symmetric matrices at the most. Let $[\mathfrak{g}]_k$ denote the grade- k part of the versor \mathfrak{g} . For example a versor \mathfrak{g} with maximal grade four can be written as $\mathfrak{g} = [\mathfrak{g}]_0 + [\mathfrak{g}]_2 + [\mathfrak{g}]_4$. The part of maximal grade $[\mathfrak{g}]_{\max}$ is always a blade, see [10]. The non-null vector $\mathfrak{v} \in \wedge^1 V$ that is contained in $\text{NO}([\mathfrak{g}]_{\max})$ satisfies

$$\mathfrak{v}\mathfrak{g}_{\max} = \mathfrak{v} \cdot \mathfrak{g}_{\max}. \quad (6)$$

Hence, $\mathfrak{g}' = \mathfrak{g}\mathfrak{v}^{-1}$ reduces the grade of the maximum blade of \mathfrak{g} by one. With remark 4 this can be simplified to $\mathfrak{g}' = \mathfrak{g}\mathfrak{v}^*$ for homogeneous Clifford algebra models. Moreover, for vectors $\mathfrak{v} \in \wedge^1 V$ we have $\mathfrak{v}^* = -\mathfrak{v}$, and therefore, we can use $\mathfrak{g}' = \mathfrak{g}\mathfrak{v}$ to reduce the degree of the initial versor \mathfrak{g} . The result of this product is again a versor. Repeated application of this process results in a set of vectors $\mathfrak{v}_1, \dots, \mathfrak{v}_{\max}$ whose geometric product results in the versor \mathfrak{g} except for a real factor. Note that this procedure may fail for null versors. We show this algorithm in an example

Example 1. *Let $K \in \text{PGL}(\mathbb{P}^3(\mathbb{R}))$ be a collineation of $\mathbb{P}^3(\mathbb{R})$ that acts on points*

$$K = \begin{pmatrix} 1 & 0 & 3 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 2 & 1 & 0 \\ 1 & 1 & 2 & 1 \end{pmatrix}.$$

To get a versor $\mathfrak{g} \in \mathcal{C}\ell_{(3,3)}^+$ corresponding to this collineation we have to solve the system (5).

$$\begin{aligned} 2(g_7+g_{17}) &= 0, & 2(g_{18}-g_3) &= 3, \\ 2(g_{19}+g_2) &= 0, & -2(g_{26}+g_{16}) &= 1, \\ 2(g_5+g_{25}) &= 0, & 2(g_6-g_{22}) &= 1, \\ 2(g_{15}-g_{30}) &= 1, & 2(g_8+g_{28}) &= 2, \\ 2(g_{21}+2g_{10}) &= 0, & -2(g_{31}+g_{14}) &= 1, \\ 2(g_{11}-g_{27}) &= 1, & 2(g_{23}+g_{12}) &= 2, \end{aligned}$$

$$\begin{aligned} g_1 - g_{20} - g_{24} - g_{32} - g_{29} + g_9 + g_4 + g_{13} &= 1, \\ g_{24} - g_9 + g_{20} - g_{13} - g_{32} + g_1 + g_4 - g_{29} &= 1, \\ g_1 - g_{13} - g_{32} - g_4 + g_{29} + g_9 - g_{24} + g_{20} &= 1, \\ g_{24} + g_{13} + g_{29} + g_1 - g_4 - g_9 - g_{20} - g_{32} &= 1. \end{aligned}$$

There are two possibilities to guarantee that the resulting versor is in the Spin group, i.e., $\mathfrak{g}\mathfrak{g}^* = 1$ or $\mathfrak{g}\mathfrak{g}^* = -1$. We compute both solutions and start with the constraint equations implied by eq. $\alpha(\mathfrak{g})\mathfrak{v}\mathfrak{g}^* \in \wedge^1 V$ for all $\mathfrak{v} \in \wedge^1 V$ and $\mathfrak{g}\mathfrak{g}^* = 1$. The corresponding Spin group element has the form:

$$\begin{aligned} \mathfrak{g}_+ &= \frac{1}{8\sqrt{2}} (7e_0 + 6e_{12} - 6e_{13} + e_{14} - 2e_{15} - 6e_{23} \\ &\quad + 6e_{24} - e_{25} - 2e_{26} + 2e_{34} + 6e_{35} - 5e_{36} - 4e_{45} \\ &\quad + 2e_{46} + 6e_{1234} - 4e_{56} + 6e_{1235} - 6e_{1236} \\ &\quad - 5e_{1245} + 2e_{1246} - 4e_{1256} + 2e_{1345} - e_{1346} \\ &\quad + 2e_{1356} - 2e_{2345} + 2e_{2346} + e_{2356} - 2e_{2456} \\ &\quad - e_{123456}). \end{aligned}$$

If we demand that $\mathfrak{g}\mathfrak{g}^* = -1$ the resulting Spin group element is computed as

$$\begin{aligned} \mathfrak{g}_- &= \frac{1}{8\sqrt{2}} (e_0 - 6e_{12} - 6e_{13} - e_{14} + 2e_{15} + 4e_{16} \\ &\quad + 6e_{23} + 2e_{24} + e_{25} + 2e_{26} + 2e_{34} + 2e_{35} + 5e_{36} \\ &\quad + 2e_{46} - 6e_{1234} + 6e_{1235} + 6e_{1236} + 5e_{1245} \\ &\quad - 2e_{1246} + 6e_{1345} + e_{1346} - 2e_{1356} - 4e_{1456} \\ &\quad - 2e_{2345} + 6e_{2346} - e_{2356} - 2e_{2456} - 4e_{3456} \\ &\quad - 7e_{123456}). \end{aligned}$$

Both elements \mathfrak{g}_+ and \mathfrak{g}_- correspond to the same collineation whose entries can be computed with the coefficients of \mathfrak{g}_+ and \mathfrak{g}_- and $(m_{i,j}), i, j = 0, \dots, 3$.

Depending on the matrix of the projective transformation, i.e., if it describes a collineation or correlation that acts on points or planes the corresponding versor can be computed with the use of the corresponding systems derived in the last section.

Remark 5. The matrix

$$K = \begin{pmatrix} -1 & 0 & 3 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 2 & 1 & 0 \\ 1 & 1 & 2 & 1 \end{pmatrix}$$

results in a versor with complex entries. Thus, we have to work over the complex numbers.

Moreover, it can be verified that $\mathfrak{g}_+ \hat{=} \mathfrak{J}\mathfrak{g}_-$, where $\hat{=}$ means is equal up to a scalar factor. Multiplication with the pseudoscalar does not change the action of the versor since

$$\begin{aligned} (\mathfrak{J}\mathfrak{g})\mathfrak{v}(\mathfrak{J}\mathfrak{g})^* &= \alpha(\mathfrak{J})\alpha(\mathfrak{g})\mathfrak{v}\mathfrak{g}^*\mathfrak{J}^* \\ &= \mathfrak{J}\mathfrak{J}^*\alpha(\mathfrak{g})\mathfrak{v}\mathfrak{g}^* \\ &\hat{=} \alpha(\mathfrak{g})\mathfrak{v}\mathfrak{g}^*. \end{aligned}$$

Remark 6. Multiplication with an element of the center $\mathcal{C}(\mathcal{C}\ell_{(3,3)}) = ae_0 + be_{123456}$, $a, b \in \mathbb{C}$ has no effect on the sandwich action of a versor. Therefore, we have a group isomorphism

$$\text{PGL}(4, \mathbb{C}) / \mathcal{C}(\text{PGL}(4, \mathbb{C})) \rightarrow \text{Pin}(\mathcal{C}\ell_{(3,3)}) / \mathcal{C}(\mathcal{C}\ell_{(3,3)})$$

Now we factorize the element \mathfrak{g}_+ into vectors that correspond to null polarities. Due to the homogeneous setting we start with the numerator of \mathfrak{g}_+ of ex. 1 denoted by $\mathfrak{g} (= 8\sqrt{2}\mathfrak{g}_+)$.

Example 2. To factorize \mathfrak{g} we first observe that $[\mathfrak{g}]_{\max} = [\mathfrak{g}]_6 = -e_{123456}$. Multiplication with a

scalar does not change the outer product null space of a k -blade, and therefore, it follows

$$\text{NO}([g]_{\max}) = \text{NO}(e_{123456}) = \{v \in \bigwedge^1 V\}.$$

We can now chose an arbitrary non-null vector $v_1 \in \text{NO}([g]_{\max})$. Let $v_1 = e_1 + e_4$ be the first factor. The versor $g_1 = gv_1$ has maximal grade five and the grade-5 part is given by

$$[g_1]_5 \hat{=} e_{23456} - 4e_{12345} + 4e_{12346} - 3e_{12456} + e_{13456}.$$

The outer product null space of $[g_1]_5$ is computed as

$$\text{NO}([g_1]_5) = \{a \in \bigwedge^1 V \mid a \wedge [g_1]_5 = 0\}.$$

This results in the set of all vectors $a = \sum_{i=1}^6 a_i e_i$ satisfying $a_1 - a_2 - 3a_3 + 4a_5 + 4a_6 = 0$. We chose the non-null vector $v_2 = 4e_2 + e_5 \in \text{NO}([g_1]_5)$ and compute the grade-four part of $g_2 = g_1 v_2$

$$[g_2]_4 \hat{=} e_{2356} - 12e_{1234} + 5e_{2345} - 4e_{1236} + 4e_{1235} \\ + e_{1245} + e_{1345} - 8e_{2346} + 4e_{2456} - e_{3456} \\ + 8e_{1246} - 3e_{1256} - 4e_{1346} + e_{1356} + e_{1456}.$$

Again we determine a vector contained in the outer product null space of $[g_2]_4$. This null space is three-dimensional and spanned by the vectors

$$b_1 = e_1 + 5e_2 + e_6, \quad b_2 = 4e_1 + 8e_2 + e_5 \\ b_3 = e_1 + e_2 + e_4, \quad b_4 = e_1 + 4e_2 - e_3.$$

We chose the non-null vector b_3 and set $v_3 = b_3$. In the next step we have to compute the outer product null space of $[g_3]_3$ with $g_3 = g_2 v_3$ that is two-dimensional and spanned by:

$$b_1 = e_1 + 4e_2 - 3e_3, \quad b_2 = 3e_1 + 7e_2 - e_4 + e_5, \\ b_3 = 4e_2 - e_4 + e_6$$

To keep the factors simple we set

$$v_4 = b_1 + b_3 = -e_1 + e_3 - e_4 + e_6.$$

The reduced versor $g_4 = g_3 v_4$ has maximal grade two and the outer product null space of the grade-two part is spanned by the two vectors:

$$b_1 = 2e_1 + 3e_2 + e_3 - e_4 + e_5,$$

$$b_2 = 4e_2 - e_4 + e_6.$$

Note that the vector b_2 is a null vector, and therefore, we chose $v_5 = b_1$ as the next factor. The last factor v_6 is obtained by

$$v_6 = g_4 v_5 \\ = 8e_1 + 8e_2 + 4e_3 - 3e_4 + 4e_5 - e_6.$$

This results in the factorization

$$v_6 v_5 v_4 v_3 v_2 v_1 \hat{=} g.$$

The next step is to carry over the factorization to projective transformations. Each vector corresponds to a null polarity. Therefore, we aim at a matrix product of the form

$$M_6 M_5 M_4 M_3 M_2 M_1 \hat{=} K.$$

Since we started with a collineation that maps points to points we have to ensure that the vectors are transferred to projective mappings of the right type. For example the product $M_2 M_1$ must be the matrix corresponding to a collineation that maps points to points. Hence, M_1 has to correspond to a null polarity acting on points and M_2 to a null polarity acting on planes. Thus, the matrix product has to be understand as

$$M_6 M_5^* M_4 M_3^* M_2 M_1^* \hat{=} K,$$

where M_i^* , $i = 1, 3, 5$ is a null polarity acting on points and M_i , $i = 2, 4, 6$ is a null polarity acting on planes. We compute the matrices M_i^* and M_i with the matrix representations computed in section 3.2. This results in the six null polarities:

$$M_1^* = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad M_2 = \begin{pmatrix} 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & -1 \\ -4 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \\ M_3^* = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 \\ 0 & -1 & 1 & 0 \end{pmatrix}, \quad M_4 = \begin{pmatrix} 0 & -1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & -1 & 0 & -1 \\ -1 & 0 & 1 & 0 \end{pmatrix}, \\ M_5^* = \begin{pmatrix} 0 & 1 & -1 & 0 \\ -1 & 0 & -1 & 3 \\ 1 & 1 & 0 & -2 \\ 0 & -3 & 2 & 0 \end{pmatrix}, \quad M_6 = \begin{pmatrix} 0 & 8 & 8 & 4 \\ -8 & 0 & -1 & -4 \\ -8 & 1 & 0 & -3 \\ -4 & 4 & 3 & 0 \end{pmatrix}.$$

With these six skew symmetric matrices it can be verified that

$$M_6 M_5^* M_4 M_3^* M_2 M_1^* = -\frac{1}{4} K.$$

Thus, we have found one possible factorization of the initial projective mapping into six null polarities.

It is worth to check the maximal grade of a versor that is solution of the other systems that were derived in section 3.2 if a factorization of a 4×4 matrix is searched. The maximal grade of the versor determines the number of skew-symmetric matrices that are necessary to factorize the given matrix and for correlations the maximal possible grade is five.

5. THE HOMOGENEOUS CLIFFORD ALGEBRA MODEL OF LIE'S SPHERE GEOMETRY

The same construction that we described for Klein's quadric can be applied to any quadric. As a second example that shall demonstrate the power of this calculus we examine Lie sphere geometry in a Clifford algebra context. Lie sphere geometry is the geometry of oriented spheres. Especially, for the three-dimensional case the set of oriented spheres can be mapped to a hyperquadric L_1^4 in five-dimensional projective space $\mathbb{P}^5(\mathbb{R})$. The construction goes back to S. LIE and was treated again by W. BLASCHKE, cf. [1]. A modern treatment of this topic can be found in [2]. Moreover, the Lie construction can be achieved for arbitrary dimension.

5.1 Lie's Quadric

A point model for the set of oriented hyperspheres, hyperplanes, and points (considered as spheres of radius 0) of \mathbb{R}^n is given by the projective hyperquadric

$$L_1^{n+1}: -x_0^2 + x_1^2 + \dots + x_{n+1}^2 - x_{n+2}^2 = 0.$$

For our purposes it is convenient that we restrict ourselves to the case of oriented spheres in

three-dimensional Euclidean space. Nevertheless, we formulate the calculus for arbitrary dimensions. The quadric $L_1^{n+1} \subset \mathbb{P}^{n+2}(\mathbb{R})$ is of dimension $n+1$, degree 2, and is called Lie's quadric. The maximal dimension of subspaces contained by L_1^{n+1} is 1, and therefore, there are no two-spaces contained entirely in L_1^{n+1} . Oriented hyperspheres, hyperplanes, and points are represented in Lie coordinates as shown in Table 1. It is not difficult to recover the Euclidean representation from Lie coordinates. If $x_0 + x_1 = 0$ and if $x_{n+2} = 0$ we have the point at infinity. If $x_{n+2} \neq 0$ we bring the point to the form $(h, -h, N, 1)^T \mathbb{R}$ by dividing by x_{n+2} . If $x_0 + x_1 \neq 0$ and if $x_{n+2} = 0$, we have a proper point. We obtain its normal form by dividing by $x_0 + x_1$. The last case is if $x_{n+2} \neq 0$. In this case we have an oriented sphere. Again we get its normal form through division by $x_0 + x_1$.

The fundamental invariant of Lie sphere geometry is the oriented contact of spheres. It is not difficult to show that two spheres are in oriented contact if, and only if, their Lie coordinates $s_1, s_2 \in L_1^{n+1}$ satisfy $\ell(s_1, s_2) = 0$, where $\ell(\cdot, \cdot)$ denotes the bilinear form corresponding to L_1^{n+1} .

Especially for $n = 3$ the lines on L_1^4 correspond to so called parabolic pencils of spheres. These pencils consist of all oriented spheres with one common point of contact. Furthermore, each parabolic pencil contains exactly one point, i.e., sphere of radius 0. If this point sphere is not ∞ the pencil contains exactly one oriented hyperplane Σ .

Remark 7. Conics on Lie's quadric correspond to Dupin cyclides, that are the envelopes of two one-parameter families of spheres.

The group of Lie transformations shows up as the group of projective automorphisms of L_1^{n+1} . This group is isomorphic to $O(n+1, 2)/\pm 1$, see CECIL [2]. Since the Pin group of the Clifford algebra $\mathcal{C}\ell_{(n+1, 2, 0)}$ is a double cover of $O(n+1, 2)$

Euclidean	Lie
points: $u \in \mathbb{R}^n$	$\left(\frac{1+u \cdot u}{2}, \frac{1-u \cdot u}{2}, u_1, \dots, u_n, 0\right)^T \mathbb{R}$
∞	$(1, -1, 0, \dots, 0, 0)^T \mathbb{R}$
sphere: center $p \in \mathbb{R}^n$, signed radius r	$\left(\frac{1+p \cdot p - r^2}{2}, \frac{1-p \cdot p + r^2}{2}, p_1, \dots, p_n, r\right)^T \mathbb{R}$
planes: $u \cdot N = h$, unit normal $N \in \mathbb{R}^n$	$(h, -h, N_1, \dots, N_n, 1)^T \mathbb{R}$

Table 1: Correspondence between Euclidean entities and Lie-coordinates.

we can use this group to describe Lie transformations.

5.2 The homogeneous Clifford Algebra Model corresponding to Lie Sphere Geometry

In this section we discuss the Clifford algebra model for Lie Sphere Geometry in the three-dimensional case. Therefore, the projective space we are dealing with is a five-dimensional space $\mathbb{P}^5(\mathbb{R})$. The homogeneous Clifford algebra model is obtained with the six-dimensional real vector space \mathbb{R}^6 as a model for the projective image space together with the quadratic form of Lie's quadric

$$Q = \begin{pmatrix} -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix}.$$

This algebra has signature $(p, q, r) = (4, 2, 0)$ and is of dimension $2^6 = 64$. Again the advantage of the Clifford algebra lies in the common description of the application of Lie transformations. Arbitrary projective subspaces of $\mathbb{P}^5(\mathbb{R})$ are transformed by the sandwich operator. As an example we determine all Lie inversions that leave the point at infinity fixed, i.e., the subgroup of *Laguerre* transformations. The point at infinity has the form $\mathfrak{p} = e_1 - e_2$, compare to Table 1. A general invertible vector is given by $\mathfrak{a} = \sum_{i=1}^6 a_i e_i$ with $\mathfrak{a}\mathfrak{a} \neq 0$. The application of the sandwich

operator to \mathfrak{p} results in

$$\begin{aligned} \alpha(\mathfrak{a})\mathfrak{p}\mathfrak{a}^* &= -2(a_1 + a_2)a_3e_3 - 2(a_1 + a_2)a_4e_4 \\ &\quad - (a_1^2 + a_2^2 + a_3^2 + a_4^2 + a_5^2 - a_6^2 + 2a_2a_1)e_1 \\ &\quad - (a_1^2 + a_2^2 - a_3^2 - a_4^2 - a_5^2 + a_6^2 + 2a_2a_1)e_2 \\ &\quad - 2(a_1 + a_2)a_5e_5 - 2(a_1 + a_2)a_6e_6. \end{aligned}$$

To guarantee that this entity represents the point at infinity, we first see that $a_1 + a_2 = 0$. With this condition the coefficients of e_3, e_4, e_5 , and e_6 vanish. Moreover, the sum of the coefficients e_1 and e_2 has to vanish. This results in

$$-2a_1^2 - 2a_2^2 - 4a_2a_1 = -2(a_1 + a_2)^2 = 0.$$

Therefore, the only condition to a Lie inversion that it represents a *Laguerre* transformation is given by $a_1 + a_2 = 0$ and the subgroup of *Laguerre* transformations is generated by all vectors with $a_1 + a_2 = 0$.

Remark 8. *Analogue to Th. 1 we can formulate a similar theorem for Lie sphere geometry in arbitrary dimensions. Since the projective model space for n-dimensional Lie sphere geometry has dimension $n + 2$, the vector space for the homogeneous Clifford algebra model has dimension $n + 3$. That means the highest grade is equal to $n + 3$, and therefore, every group element can be written as the composition of $n + 3$ vectors at the most. Especially for the case $n = 3$, we have similar results as for Klein's quadric. In this case six involutions are necessary to generate the whole group.*

Let us reformulate this remark as theorem.

Theorem 2. *Every Lie transformation in n -dimensional space is the composition of $n+3$ involutions, that correspond to the sandwich action of vectors.*

6. CONCLUSION

With the help of the homogeneous Clifford algebra model corresponding to line geometry we introduced null polarities as fundamental involutions. After we transferred linear line manifolds to this model, we derived the correspondence between regular projective transformations and versors. Every versor can be factorized into vectors that correspond to null polarities. Therefore, every 4×4 matrix corresponding to a projective transformation can be expressed as the product of six skew symmetric matrices at the most. Moreover, we presented a homogeneous Clifford algebra model corresponding to Lie sphere geometry.

ACKNOWLEDGMENTS

This work was supported by the research project "Line Geometry for Lightweight Structures", funded by the DFG (German Research Foundation) as part of the SPP 1542.

REFERENCES

- [1] Blaschke, W. Thomsen, G. Vorlesungen über Differentialgeometrie und Geometrische Grundlagen von Einsteins Relativitätstheorie Bd.3: Differentialgeometrie der Kreise und Kugeln. Springer, Berlin, 1929.
- [2] Cecil, T.E. Lie Sphere Geometry, 2 ed. Springer, New York, 2008
- [3] Fontijne, D. Efficient Implementation of Geometric Algebra. PhD thesis, University of Amsterdam, 2007.
- [4] Gallier, J. Clifford Algebras, Clifford Groups, and a Generalization of the Quaternions: The Pin and Spin Groups. 2008, arxiv.org, available at <http://arxiv.org/abs/0805.0311>.
- [5] Garling, D.J.H. Clifford Algebras: An Introduction. Camb. Uni. Press, 2011.
- [6] Gunn, C. Geometry, Kinematics, and Rigid Body Mechanics in Cayley-Klein Geometries. PhD Thesis, TU Berlin, 2011.
- [7] Hestenes, D. Sobczyk, G. Clifford Algebra to Geometric Calculus: A unified Language for Mathematics and Physics. D. Reidel, Dordrecht and Boston, 1984
- [8] Klawitter, D. A Clifford algebraic Approach to Line Geometry. Advances in Applied Clifford Algebras, doi: 10.1007/s00006-014-0459-z, 2014.
- [9] Klawitter, D. Clifford Algebras - Geometric Modelling and Chain Geometries with Application in Kinematics. PhD thesis, Technical University of Dresden, 2014.
- [10] Perwass, C.B.U. Geometric Algebra with Applications in Engineering. Springer, Berlin and Heidelberg, 2009.
- [11] Porteous, Ian R. Clifford Algebras and the Classical Groups. Cambridge University Press, 1995.
- [12] Pottmann, H. Wallner, J. Computational Line Geometry. Springer, Berlin and New York, 2001
- [13] Weiss, G. Vorlesungen aus Liniengeometrie. Unpublished manuscript, Institute of Geometry, Vienna University of Technology, 1991.
- [14] Selig, J.M. Geometric Fundamentals of Robotics. Springer, 2005.

ABOUT THE AUTHOR

1. Daniel Klawitter, Dipl.-Math., is a PhD student at the Institute of Geometry at Dresden University of Technology, Germany. His research interests are Clifford algebras, line geometry, computational and theoretical kinematics, differential geometry and computer aided geometric design. He can be reached by e-mail: daniel.klawitter@tu-dresden.de or through postal address: TU Dresden, Institute of Geometry, 01062 Dresden, Germany.

AN OPEN SOURCE WORK FLOW: A DIGITAL BUILDING RECONSTRUCTION AND ITS PRESENTATION

Alessio MAZZUCCHI
University of Trento, Italy

ABSTRACT:

Sharing thoughts and ideas is one of the best methods to spur innovation, this is why the Open Source world captures the interest of many people, including the author of this work. It is generally believed that it is hard to achieve good and productive results using only Open Source tools. Hence, the need to prove the contrary. The right opportunity came when the author was assigned a digital reconstruction of an old Asburgic building. The architecture hosts today the Central Postal Office of Trento, and it shows no sign of its past appearance. Bring back the Asburgic look of the edifice, using only Open Source programs, seemed to be the perfect challenge. Keep on reading to know how this story ends...

Keywords: Open Source, CGI, Animation, Modelling, Blender

1. INTRODUCTION

Almost every proprietary professional program has its own equivalent Open Source¹ version. There is a general belief that OS programs are unable to complete the most advanced challenges that one faces in his working life. But, is it really so? Are OS softwares really unsuited to operate in highly demanding tasks? There is no simple answer to this question. This paper wants to prove that, when it comes to architectural animations and presentation, the OS way is a real alternative.

In the architectural world, the average user finds it difficult to understand the feasibility of developing a project of 3d visualization using only OS programs. This for two main false beliefs: the first one being that with OS one can not reach a good level of quality and secondly that all major studios use proprietary programs. Thus, for the sake of one's career, it is better to learn proprietary programs². This paper will try to prove the first point wrong by giving an

idea of the possibilities offered by OS softwares. The first aim of this work is to test the potentialities of OS programs. Secondly it hopes to do its part in spreading the knowledge on OS softwares. Since not all the available OS programs have reached the same level of development, the conclusions of this paper will refer only to the programs used to carry on the project. Before talking about the conclusion, it is better to start by saying where did the inspiration come from. Everything began when the author had to create a digital reconstruction of an ancient building, currently replaced by a new one. This consignment was part of the "*Disegno Automatico*" exam of the course of Architecture and Building Engineering³ of the University of Trento, in Italy. The terms of the commission were the followings: produce a digital animation and a presentation, to show the various stages of the work. This looked like the perfect opportunity to test if OS programs can be used to carry on a project, from the start to the end. The only constraint: the use of OS tools. After reflection, Blender has been chosen

¹From now onwards OS.

²This second point will not be addressed by this paper, but it could be the subject of a future work.

³The author is enrolled in this Master's degree.

for the 3d and for the compositing tasks. Gimp and Inkscape were employed for image manipulation, while the presentation was created with the HTML language, Javascript and Mootools framework⁴. All programs use Ubuntu Linux distro as their operative system.

Both the animation and the presentation have been completed using only OS tools. The animation is quite realistic; it combines a 3d model with a real video. The layout of the presentation is clear and the final result is similar to what expected. Even though the results overcome the expectations, improvements can still be made. After this short overview, the next section will introduce the different phases of the project.

2. WORK PHASES

As previously mentioned, the building to digitally model was “*Palazzo Poste*”, a historic building of Trento, Italy which hosts today the local central siege of the Italian Mail Company⁵. Between 1929 and 1934 it underwent drastic renovations, which completely distort its prior appearance. The target of this project was to revive the historic look of the building.

As every novel is composed by chapters, so it is this work. This paragraph will point out the phases of the project, which will be explained in depth in the following sections. Firstly, an investigation was accomplished to discover which of the available free tools, would be the most suited for the purpose of this work. Secondly, to better understand the appearance of the building, a brief historical research was conducted. Thirdly, the 3d model was realized. Fourthly, a clip of the current appearance of the Postal Office has been recorded and matched with the digital reconstruction. Some digital effects were added to create a transition between the new and the historic building. The final step was the development of the HTML presentation which was

carried out in parallel with the above mentioned tasks.

To sum up, the work can be divided into two main phases that group together the previously listed steps: the core of the work is the short video and the second part is the interactive presentation. To start, one needs the right tools and the following paragraph will present this period of choices.

2.1 Choices

The aims of the project could be fulfilled only with the correct tools. Then came the time to find, among the existing ones, the right OS programs. Blender, being a benchmark in the OS panorama, has been selected for the animations. Additionally, the Blender Foundation⁶ continuously ensures the development of new features and the Blender’s forum is really helpful, in it one can always find the answer to one’s problems. Moreover, the Blender Foundation was carrying on “The Mango Project⁷”, a CGI⁸ short movie which entails the introduction of a new internal tool for camera tracking.

The choice of the best suited render engine for this work has not been easy. After careful consideration the new Blender internal engine Cycles was selected, despite the fact that it was still in development. Cycles has been chosen because it ensures GPU rendering and the node integration in the same work ambient of Blender, a feature that other third parties engines do not have. For the presentation the idea was to realize a simple layout with easy animations. The choice of using HTML with a framework for Javascript was dictated by the necessity of having fewer constraints when creating the animation and their interactions. Other standard presentation programs like Impress are equally performing, but they do not allow the same degree of liberty. Once the tools

⁴A deeper discussion on this matter will follow in the next paragraph.

⁵Piazza Vittoria Alessandro n.1, 38122, Trento, ITALY - GPS 46.066910, 11.123755.

⁶The Blender Foundation is a non-profit organization responsible for the development of Blender. [1].

⁷Here, one can find more information about this project: [http://mango.blender.org/about/\[2\]](http://mango.blender.org/about/[2]).

⁸Computer-generated imagery.

chosen, the following step was to understand the characteristics of the building during of the Asburgic period.

2.2 Historical research



Figure 1: An historic photo of the Asburgic Building [5].

The next indispensable stage was to gather information about the building during the Asburgic era. Professor Campolongo⁹ has already studied the story of the building and he made available some old construction plans and old black-and-white photos. He found this material during his personal researches and he also indicated some historical manuscripts on the building¹⁰. The result was a set of heterogeneous data which often contradict themselves. The plans did not match with the photos. In view of these discrepancies, the author chose to base the structure of the building on the plans, while the details are based on the photos. He took the liberty of making a biased choice because the purpose of the work was to create the workflow. The historical reconstruction was just a pretext to show the strengths and the weaknesses of the programs. The final model

⁹Fabio Campolongo is teaching professor at the University of Trento and a member of the “*Soprintendenza per i Beni architettonici della Provincia autonoma di Trento*”.

¹⁰In the photo library of Trento there are several photos dated from 1922 to 1935 [5], Figure 1 is one of them and some plans dated 1988 of Mr. Setz, the architect who designed the Asburgic building [6]

does not pretend to be historically accurate, but it gives a realistic idea of how the building may have looked like during the Asburgic era. Having all the necessary information on the building, the modelling phase could be started off.

2.3 Modeling

The Asburgic building was created in Blender by using the historical plans as the principal source and the old photos as the reference for the details. The different parts of the building, with the exception of the cornices, have been modelled using the polygonal technique. The cornices were reproduced using NURBS, which allows exploiting the “extrusion along path” technique which is the fastest way to extrude a profile along a path. The Duplication Group was used to reproduce the repetitive elements of the façade. This command permits to link n-instances to an original object. This technique is similar to the function of CAD blocks, but with a different implementation.

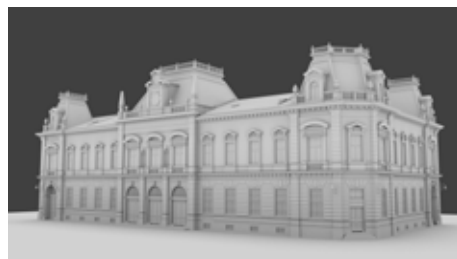


Figure 2: Clay render of the model.

2.4 Procedural material

The lack of information about real colors and materials of the building led to the choice of procedural materials as the technique to use to recreate the aspect of the surfaces. This method is more demanding than creating textures from real photos, but it grants greater liberty. Starting from simple textures, one can create complex surfaces and mimic real materials. To make a procedural material one starts by creating single procedural textures which can be added, subtracted,



Figure 3: Rendering of the rusty iron procedural material.

scaled or subjected to others transformations in the Node Editor of Blender [See Figure 4]. These textures allow the control of all parameters of the shader and the displacement of the material. At the end of this phase, the virtual reproduction of the Postal Office was completed, but its surrounding environment was not. The next step has been the creation of a scenography to contextualize the building.

2.5 Tracking

The news that the Blender Foundation was developing a new tracking tool came as a relief, as without it, the modelling of the environment would have been a time-consuming task. This new feature allowed the insertion of a digitally reconstructed object into an actual footage, right what this project needed. From the first test to the final video, the Blender tool has been developed a lot, facilitating and automating the work of the author.

The procedure followed is quite simple: the shooting of the nowadays building has been tracked and then composited together with the Asburgic one. Many tests have been carried out

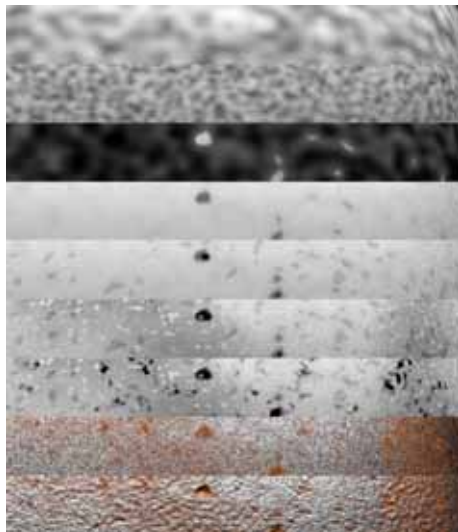


Figure 4: Different phases of the evolution of a material.

to understand how to track the video, how to set the camera, how to capture the perspective of the edifice and how to find the right points to track. A Canon 7D, with 17-55mm lens at 25 fps was used to record the video. To smooth the development of the story, the author chose to add actions to the video. The details of this choice will be explained in the next paragraph.

2.6 Physics simulation

After tracking the video, the author looked for a method to distract the audience while the nowadays buildings changes into the Aburgic one. A good magician always needs a little bit of smoke to hide his tricks. In this case the only difference lays in the amount of smoke employed: it had to cover the entire building. Blender has the right tool to fulfil this task: a particle system that creates physical simulations, which includes also the generation of smoke. First of all, it is necessary to set a domain in which the simulation can take place and to position a particles generator. Then one has to fix the parameters to control the behav-

ior of the smoke during the simulation. To clean the scene, a wind generator has to be positioned and activated to blow away the smoke at the right moment. From the simulation one obtains the voxel data of the position of the smoke for each frame, which tells to the shader where to create the smoke. At this stage, a problem aroused: Cycles did not have a reliable volume shader to render the smoke. As an alternative, the “Blender Internal” engine was used. It became necessary to create alpha masks to composite the renders obtained from Cycles and from Blender. At this stage, the author disposed of all the essential elements to ultimate the video, they just needed to be putted together.

2.7 Compositing

Blender offers two alternatives ways to make compositing: one through the “video editor” and the other through the “node editor”. Although less intuitive¹¹, the latter was chosen because its

¹¹The “node editor” lacks the timeline, tool available in the “video editor”.



Figure 5: Above the render of the smoke composed with the real footage.

flexible nature allows greater compositing freedom.

The compositing work consisted in matching the sequences of the real video, with one of the Asburgic building and with one of the smoke. The building was surrounded by many objects, cars, peoples, signals. They needed to be masked because they interfered with the rendering of the Postal Office. To create the masks for the bigger objects rotoscoping mixed with the tracking technique was used, while for the little ones, like the leaves of trees, chroma and luminance key were chosen. As the work evolved, the author got better acquainted with the tools and, consequently, the scenario of the story became clearer.



Figure 6: Above, the screen shot of the masking procedure and its result.

2.8 Plot

At the beginning of the work, the plot was written, but since tales are not fixed, this one has been adapted to the evolution of the project. The discovery of new tools and a better understanding of the existing ones provided new possible

developments of the plot. In particular, the use of camera tracking led to the decision of matching the real video with the digital version of the building. The idea was to create a transition from the nowadays edifice to the historic one and there is no better way to that than with a distraction. Consequently, it was necessary to further exploit the possibilities offered by Blender and to use physics simulation to create the smoke, to dissimulate the transition¹².

The story¹³ tells about a man, who has hallucinations, he sees a strange sphere of light fluctuating in the sky. The sphere, while passing near the Postal Office, gives the start to a smoke leak that surrounds the entire edifice. When the smoke clears away, the building is not the same: it has regained its Asburgic appearance. As in every fairy tale, the magic lasts only for a few moments, and afterwards everything falls back to normality. The work seemed to be over; on the contrary a big part of it was missing: the presentation was still a sketch.



Figure 7: Frames of the animation.

2.9 Presentation

When one thinks about the possible shapes of HTML web pages, the first figure that comes to one's mind is a rectangle: all the elements of a web page can be reduced to a rectangle. Building on this assumption, the layout of this work's

¹²The compositing task and the physics simulation have already been discussed above.

¹³The story is accompanied by an original soundtrack composed by the Italian group Outopsya.

presentation is based on this polygon. All the animations are interactions between different levels of hierarchy of rectangles. Each rectangle contains the explanation of a phase of the project and it is up to the user to decide whether or not to open a chapter. The presentation tells the story of the work, from the choice of the programs to the final video. Initially, the aim was to make the presentation functional on every browser. Due to the use of custom Javascript code this objective has not been achieved, so only Chrome supports it.



Figure 8: Homepage of the presentation.

The main obstacle to overcome during this phase has been the lack of knowledge of the author in the domain of web design. For example, the author was not aware of that, for security reasons, the communication between two HTML web pages is restricted. Luckily, the online community knew the answer and this problem was overcome thanks to an escamotage¹⁴. Despite what mentioned above, the result is an interactive presentation which fulfils the prefixed aims in terms of appearance and functionality. Nevertheless, it suffers from several bugs, e.g. the shortcuts are not always synchronized with the right section. It is now time to draw the necessary conclusions on the accomplished work.

¹⁴Here is the link where the solution was found: <http://mootools.net/forge/p/crossdomainfragment> [7]



Figure 9: A section of the presentationl

3. CONCLUSION

The practical results of this work are essentially two: a short video¹⁵ and a presentation. Both of them have been successfully completed using only OS tools. The primary objective of the project has been achieved: OS tools can perform “professional” tasks in the domain of architectural visualization.

When it comes to CGI workflow, Blender demonstrates once again¹⁶ its potentialities in every CGI aspect. The modeling tools have not shown any downside, on the contrary they have been up to standards. The creation of procedural materials has been both a challenging and a rewarding task: it required a big effort, but at the end the materials look realistic.

Even though the smoke does not have a natural aspect, the physics of the movement is credible and it looks real. Due to the lack of precision during the shooting and the inexperience of the author in this field, the tracking phase is where most of the problems arose. As a result, the digital parts of the video are not always correctly aligned. The misalignments are ascribable to the mistakes made during the recording of the video. On the whole, the final video has fulfilled the ex-

pectation, even if several small problems reduce its overall.

Also the HTML presentation has achieved the objective, since it has been created only with OS tools. The animations and the layout look as planned. The two main unsolved issues are: the presence of bugs which requests a reboot of the presentation and the fact that the animations are not always smooth. Throughout the work, the author realized that the HTML language probably has not been the best choice to carry on this task. But, having to use only OS programs, the choice of tools was limited and so the result is the best one can achieve respecting the constraint and considering the abilities of the author. Nevertheless, Internet is full of web pages with great layouts and outstanding animations that show the potentialities of the HTML5 standards.

To conclude, this work demonstrates that OS programs are suited to create architectural animations. In fact, most of the limits of this work are due to the inexperience of the author, and not to the limits of the tools¹⁷. Additionally, even if in a humble and quiet way, this paper does its part in spreading the knowledge on OS programs. But the work does not ends here, a comparison between OS and proprietary programs could be the interesting subject of an upcoming paper. In any case, the author will continue using OS programs and sustaining the efforts of the OS community.

ACKNOWLEDGMENTS

I would like to express my deep gratitude to Alessia Giuliani who helped me throughout the writing process. I would like to thank the band Outopsya for the original soundtrack of the video, Nicolò Veronesi for the help he gave me with the footage, the Blender Foundation for the great tool they developed and for the upcoming ones and Professor Fabio Campolongo for the historical material. I am also grateful to Professor Giovanna Massari who persuaded me to write this

¹⁵Here is the link of the video: https://www.youtube.com/watch?v=s_gRn4wNXH0 [4] and the link to download the finished model of the building <http://www.blendswap.com/blends/view/68769> [3].

¹⁶The official shorts of Blender Foundation already proved Blender’s capabilities.

¹⁷Please consider that the assignment is developed by someone who, at the beginning of the paper, had a medium knowledge of 3d tasks and a very low one of web design.



Figure 10: Rendering of the façade.

paper. Finally, my gratitude goes to the organisers of ICGG for the amazing opportunity they granted me.

REFERENCES

- [1] Blender Foundation. Blender foundation website, 2014. URL <http://www.blender.org/foundation/>. [Online; accessed 28-May-2014].
- [2] Blender Foundation. Project mango website, 2014. URL <http://mango.blender.org/about/>. [Online; accessed 28-May-2014].
- [3] a. mazzucchi. 3d model of the palazzo poste in the asburgic version, 2014. URL <http://www.blendswap.com/blends/view/68769>. [Online; accessed 28-May-2014].
- [4] a. mazzucchi. Poste, 2014. URL https://www.youtube.com/watch?v=s_gRn4wNXH0. [Online; accessed 28-May-2014].
- [5] S. Perdomi. Trento, piazza A. Vittoria, palazzo delle Poste. Fototeca della Provincia Autonoma di Trento, from 1922 to 1933. [Inventory name: 85893 - 85894 - 85895 - 103254 - 103255].
- [6] F. Setz. AMTS GEBÄUDE FÜR DIE K.K. POST-u. TELEGRAPHEN ANSTALT IN TRIENT. Archivio Storico della Biblioteca Comunale di Trento, 1888. [Inventory name: serie ACT3.24_1889_007].
- [7] L. Stanco. Cross-domain communication between an iframe and its parent., 2014. URL <http://mootools.net/forge/p/crossdomainfragment>. [Online; accessed 28-May-2014].

ABOUT THE AUTHORS

Mazzucchi alessio is a Student of the course of Architecture and Building Engineering of the University of Trento, in Italy. Now he is working on his master's thesis on Parametric Design.

OPTIMISATION OF BUNDLED ROUTES

Victor Parque, Masakazu Kobayashi and Masatake Higashi
Toyota Technological Institute, Japan

ABSTRACT: Designing edge bundles is computationally complex. We use a simple generative-representation through optimization to design edge bundles for complex and diverse connectivity environments. The unique point of our contribution is that not only the structure, but also the modularity of the bundle structure emerges as a result of executing a generative tree. We performed experiments using different graph architectures and discuss our encouraging results.

Keywords: bipartite graph, edge bundling, bundled route, optimisation

1. INTRODUCTION

We aim at tackling the problem on how to bundle the edges for undirected bipartite graphs, which represents the pathways for signal transmission among the components of any interconnected system. This problem is ubiquitous and significant in fields related to network optimisation, e.g. optimal wiring in neural networks, design of the optimal pathways for communication systems including transportation, optimal cable routing in circuit design and edge bundling in graph visualization.

Related research on route bundling come from Neuroscience and Graph Theory. In Neuroscience, the main results link route bundling with component placement optimization and the optimality of wiring in neural networks. In Graph Theory, the more closely related research has used the concept of convex optimization, continuous forms of clustering and diverse methods to assemble the edges in a given graph[8–13]. Despite the interesting results in both fields, still the current algorithms rely on tuning parameters, and the quality of the resulting bundles are judged qualitatively, which is in line and relevant with the visualization-approach to route bundling. In engineering applications, such as wire harnessing, one is more interested in quantitative metrics of route bundles.

The main goal of our study is how to build

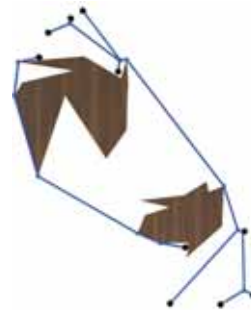


Figure 1: Modularity formation in edge bundling.

edge bundles in order to provide users with simplified graph structures and compact representations. In the majority of cases, edge bundling is reduced to the problem of building trees that minimise the global length and connectivity efforts among the nodes. In fact, edge bundling is related to the Minimum Steiner Tree, with the only difference that edge bundling permits the formation of modules of trees. Fig. 1 shows an example of such modularity formation identified by our algorithm. Since the aim is to attain shorter and compact tree structures, the formation of modules is important in situations where complete (total) edge bundling of the graph is impractical and tends to render longer structures.

Different from the previous research views, our approach focuses on the efficient and the non-parametric generation of bundled routes, aiming at optimality under the presence of arbitrary obstacles, which has remained elusive in the literature. The basic idea of our proposal is to search over the bundle pairs of similar edges by using generative tree structures in order to optimise the connectivity length.

The rest of this paper is organized as follows. Section 2 introduces the related research to our work. Section 3 introduces our approach to edge bundling. Section 4 summarizes the computational experiments we conducted in our study; and Section 5 concludes the paper.

2. RELATED WORK

The problem of bundling routes is closely related to past research concerned with optimal wiring and connectivity in graphs and networks. Particularly, since understanding how the networks form and function in the nervous system is of special interest in the neuroscience field, several studies have tackled related problems with route bundling. For example, in a seminal work, neural formation is explained as component optimization in which the save wire principle predicts that placement of sensory neurons that optimize the connectivity layout of a neural structure was proposed by Cherniak[2, 3] and a function to estimate the wire length in the layout of neural pathways was proposed in a subsequent work[1]. Additional surrogates as estimators of wiring connectivity cost was also proposed by Cuntz et al. [4, 5]. Since it is not clear what class of algorithms give birth to the wire connectivity in neural networks, Clune et al. [6] proposed that wiring and modularity formation are byproducts of evolving neural networks with both performance and connectivity costs. However, the use of heuristics such as genetic algorithms induces variability and inaccuracy due to inherent randomness to escape local optima.

Route bundling has been studied in the past decade for graph visualization. A detailed re-

view on the accomplishments is provided by Zhou et al.[7]. Basically, given a graph, the problem lies in deciding *what* to group, *when* to group and *how* to group. The *what* question relates to whether it should be nodes or edges the elements being merged, the *when* question relates to the time sequence for the merging elements, and the *how* question relates to the geometric procedures to map the merging elements into a new and single element. Succinctly, Phan et al. proposed an algorithm to minimize edge crossing (*what*) while modifying the position of nodes in a tree by using hierarchical clustering (*when*) and bounding boxes (*how*)[8]. Holten and VanWijk (2009)[9] proposed an algorithm based on attracting forces (*how-when*) among edges (*what*) where the bundling level is determined by user choice; Ganser et al. (2011)[10] proposed to use kd-trees (*when*) to cluster the routes (*what*) continuously and minimise ink usage (*how*) while merging; Ersoy et al. (2011)[11] proposed the skeleton-based with fast marching method (*how*) to ensemble the image-based routes (*what*) under adaptive parametric cuts of hierarchical clustering to control convergence to optima (*when*); and more recently Hurter et al. (2012)[12] proposed the kernel-based (*how-when*) grouping of routes (*what*). However, the performance of the above algorithms are dependant on the pre-tuning certain parameters: window size[8], edge stiffness constant[9], maximum turning point[10], edge similarity threshold[11] and kernel function[12].

We propose a generative approach to build edge bundles given information of a bipartite graph and a polygonal set, which represents the set of obstacles in the map. The basic idea is to search over the bounded space of similar routes by using generative tree structures to optimise the connectivity length. The unique point of our approach is that it aims at generating not only the optimal bundles, but also the independent modules considering an efficient and parameter-free constrained optimization.

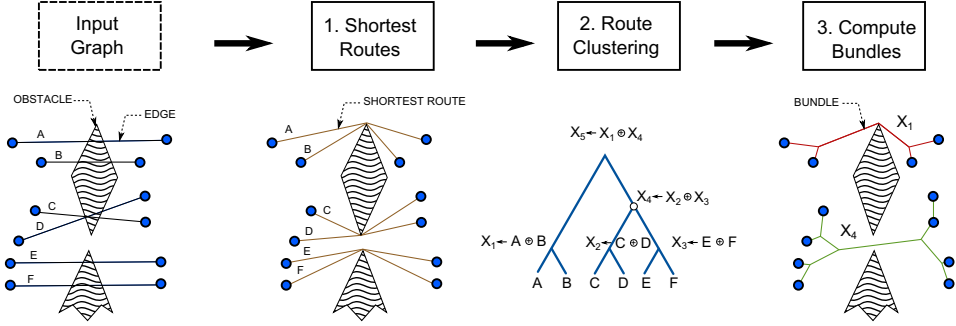


Figure 2: Basic idea and the main steps in our proposed method for edge bundling

3. OUR APPROACH

To give a general view of our algorithm, Fig. 2 shows a basic example of the main steps of our approach, where the main inputs and outputs are summarized by Algorithm 1.

Algorithm 1: Route bundling

Input : Obstacle Set \mathcal{P}
 Undirected Graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$

Output : Bundle Set \mathcal{B}

- 1 $\mathcal{R} \leftarrow \text{Find shortest routes}(\mathcal{G}, \mathcal{P})$
 - 2 $\mathcal{T} \leftarrow \text{Cluster hierarchically}(\mathcal{R})$
 - 3 $\mathcal{B} \leftarrow \text{Compute bundles}(\mathcal{T}, \mathcal{P})$
-

The inputs for Algorithm 1 are the following:

- An obstacle set \mathcal{P} , which is a collection of polygons, where every polygon denotes the constraints (holes) to avoid in the given space. Fig. 2 shows an example of set of obstacles consisting of two polygons.
- A bipartite graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with $|\mathcal{V}| = n$ nodes and $|\mathcal{E}| = m$ edges, where every node denotes either a source or target for the connectivity between the coordinates of two elements in a given space. Fig. 2 shows an example of a bipartite graph with six edges. Note that every coordinate of the vertices of the graph \mathcal{G} is outside the set of polygons \mathcal{P} .

The output for Algorithm 1 is the bundle set \mathcal{B} containing the collection of assembled edges of the bipartite graph \mathcal{G} that avoid the given obstacle constraints \mathcal{P} .

For simplicity and without loss of generality, we assume that the coordinates of the vertices of the bipartite graph, the obstacle set, and the bundle set are in \mathbb{R}^2 . The extension to higher dimensions is straightforward

Succinctly, the main steps in our algorithm are summarized as follows:

- First, we find the set of shortest paths \mathcal{R} for every edge in \mathcal{G} .
- Second, we find the tree \mathcal{T} which represents the similarity structure of the set of shortest paths \mathcal{R} .
- Third, we process every non-terminal node X_i in \mathcal{T} selectively: bundle the routes if the ensemble is shorter.

The following subsections will discuss details of the procedures outlined above.

3.1 Generating shortest routes

In order to find an initial set of feasible paths that avoid the given set of obstacles \mathcal{P} , we first compute the shortest paths between opposite vertices of every edge e_u in \mathcal{G} , where $u \in [m]$, by using



Figure 3: Example of shortest routes

using Algorithm 2, wherein the *ShortPath* procedure is a path-finding algorithm. We use an implementation of the A* algorithm¹, considering:

- a visibility graph (mesh) constructed with the outer boundaries of the obstacles, enabling the fewer number of vertices in the mesh and the faster navigation avoiding the obstacles, and
- a heuristic cost equivalent to the straight Euclidean distance to the goal.

Algorithm 2: Find shortest routes

Input : Undirected Graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$,
Obstacle Set \mathcal{P}

Output : Bundle Set \mathcal{R}

```

1  $\mathcal{R} \leftarrow \emptyset$ 
2 for each edge  $e_u$  in  $\mathcal{G}$  do
3    $r_u \leftarrow \text{ShortPath}(e_u, \mathcal{P})$ 
4    $\mathcal{R} \leftarrow \mathcal{R} \cup \{r_u\}$ 
5 return  $\mathcal{R}$ 

```

The result in Algorithm 2 is a collection of shortest paths \mathcal{R} , where every subscript $r_u \in \mathcal{R}$ denotes the sequence of points that belong to the shortest path of e_u . Fig. 3 shows an example

¹in Java v.1.6, <https://code.google.com/p/straightedge/>

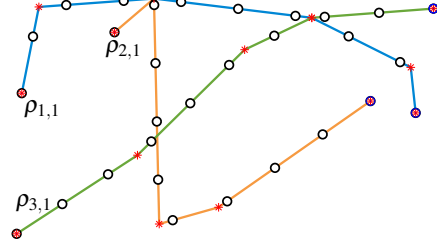


Figure 4: Example of interpolation for $N = 10$

of the shortest routes for a bipartite graph with seven edges in a given map with two obstacles.

3.2 Clustering

Since the shortest paths of the edges represent the initial solutions to compute the topology of the bundled routes, we need an explicit sequence (order) to assemble the routes. Intuitively, similar routes should be assembled before dissimilar ones: the assembly order can thus be easily encoded in a hierarchical structure, i.e., a tree \mathcal{T} in an hierarchical clustering algorithm, in which the terminal nodes (*leaves*) denote the elements of the set of routes \mathcal{R} , and the non-terminal nodes (*parents*) denote the bundles generated by assembling the leaf nodes.

Note that, in terms of formal language theory, the non-terminal nodes of the tree \mathcal{T} also represent the *grammar* to build a finite set of ensembles. Furthermore, the tree encodes relative distances too: *sibling* nodes are structurally more *similar* than to any other leaf, where the measure of similarity (distance) is computed as follows:

$$d(r_u, r_v) = \sqrt{\sum_{k=1}^N \|\rho_{u,k} - \rho_{v,k}\|^2}, \quad (1)$$

where $d(r_u, r_v)$ represents the distance between the routes r_u and r_v , N represents the number of equidistant sampled (*interpolated*) points in both routes, and $\rho_{u,k}$ and $\rho_{v,k}$ are the k -th corresponding sampled points in the route r_u and r_v , respectively. An example on how ten equidistant

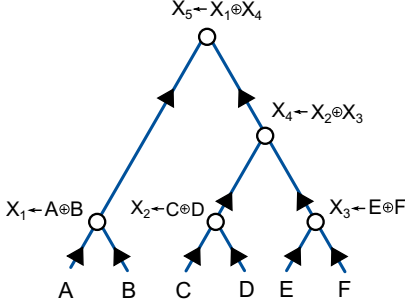


Figure 5: Ordering of bundling using the tree \mathcal{T}

points (circles) are sampled along three arbitrary routes is shown by Fig. 4. By using the above distance relation, we have the advantage to compute the similarity degree of two routes even if they have different number of nodes as shown by the asterisks in red color of Fig. 4.

In order to generate the tree structure \mathcal{T} , we used the hierarchical clustering algorithm implemented in SciPy² considering the *complete/max/farthest* point linkage, which generated improved topologies compared to the single, average, weighted, centroid and median linkage methods. Devising an improved clustering algorithm for the edge bundling problem is left for future work. An example of a hierarchical tree is shown by Fig. 5.

3.3 Bundling

Since the tree structure \mathcal{T} represents the sequence to assemble the routes \mathcal{R} , the construction of bundles is a hierarchy: by pairs and in bottom-up order, and following the *hierarchical labeling* of the tree \mathcal{T} , where parent nodes are labeled with ordinal integers after the child nodes are labeled as well. In this scheme, the root of the tree \mathcal{T} has the highest label among all labels. Particularly, an example of hierarchical labeling is shown by the subscripts of the X in Fig. 5.

Structurally, a **bundle** X_i is a non-terminal, hierarchically labeled, node of \mathcal{T} encoded by the

²Python 3.3.2, 64 bit

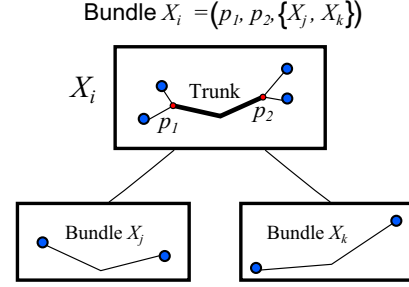


Figure 6: Encoding of a bundle X_i

tuple (p_1, p_2, L) , where p_1 and p_2 are bifurcation points and L is the set of *leaves* of the bundle X_i as shown by Fig. 6. The *trunk* of X_i is the shortest route between p_1 and p_2 . In this context, the bundle X_i implicitly represents the instructions in the form: $X_i \leftarrow X_j \oplus X_k$, where X_j and X_k are the *leaves* of X_i in \mathcal{T} , and the \oplus operator denotes the *Build bundle* function used in Algorithm 3, and described by Algorithm 4.

Algorithm 3: Compute Bundles

Input : Tree \mathcal{T} and Obstacle Set \mathcal{P}

Output : Bundle Set \mathcal{B}

```

1  $\mathcal{B} \leftarrow \emptyset$ 
2 for each hierarchically-labeled  $X_i$  in  $\mathcal{T}$  do
3    $X_i \leftarrow \text{Build bundle}(X_j, X_k, \mathcal{P})$ 
4    $\mathcal{B} \leftarrow \mathcal{B} \cup \{X_i\}$ 
5 return  $\mathcal{B}$ 

```

A key function of Algorithm 4 is to build a new bundle X_i from the input bundles X_j and X_k , which implies three roles:

- **Triangulation**: to construct a searchable space for X_i given the obstacles \mathcal{P} and the leaves X_j, X_k ,
- **Bifurcation**: to find the bifurcation points p_1 and p_2 and encode along with the leaves X_j and X_k to construct the bundle X_i , and

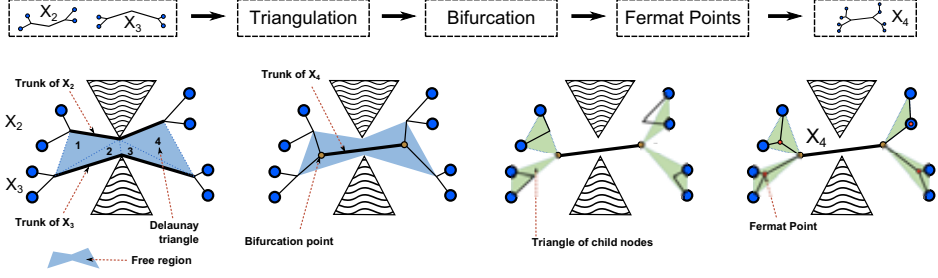


Figure 7: Example of a bundling procedure

- **Fermat Points:** to update the leaves X_j and X_k due to the insertion of the new points p_1 and p_2 into the bundle X_i .

The following paragraphs will describe these roles and the implementation aspects of Algorithm 4. For clarity of our explanation, an example of on how to build the bundle $X_4 \leftarrow X_2 \oplus X_3$ (drawn from Fig. 2) is represented by Fig. 7.

Algorithm 4: Bundle operation

Input : Bundle X_j, X_k , Obstacle Set \mathcal{P}
Output : Bundle X_i : $X_i \leftarrow X_j \oplus X_k$

- 1 $C_{j,k} \leftarrow \text{ConvexHull}(\text{Trunk}(X_j) \cup \text{Trunk}(X_k))$
- 2 $\mathcal{T} \leftarrow \text{Triangulate}(C_{j,k} - \mathcal{P})$
- 3 $(p_1, p_2) \leftarrow \underset{(p_1, p_2)}{\text{Minimize}} J(p_1, p_2, X_j, X_k, \mathcal{T})$
- 4 $X_i \leftarrow (p_1, p_2, \{X_j, X_k\})$
- 5 $X_i \leftarrow \text{Fermat Points}(X_i)$
- 6 **if** $\text{Length}(X_i) > \text{Length}(X_j) + \text{Length}(X_k)$
then
- 7 $X_i \leftarrow \emptyset$
- 8 **return** c

Triangulation (Fig. 7, 1st column). In order to find the bifurcation points, we first need to construct a *searchable region* being free-of-obstacles, for which we first delineate the convex hull $C_{j,k}$ between the *trunk* of the input bundles X_j and X_k , and then compute the difference

between the convex hull $C_{j,k}$ and the obstacle set \mathcal{P} , since the difference represents the free-of-obstacle region, wherein we aim at finding p_1 and p_2 . In most cases, the difference represents an irregular and non-convex polygonal set, thus making the search process particularly inefficient due to collision detection checks. Hence, we triangulate the difference into a *searchable region* \mathcal{T} and encode any point $p \in \mathcal{T}$, including p_1 and p_2 , using a *triangular system*: a tuple (t, r_1, r_2) , where t represents the ordinal number of the triangle in which p is: $t \in [1, |\mathcal{T}|]$, and $r_1, r_2 \in [0, 1]$. There exists a simple mapping[13] that allows the bijection of this tuple into Cartesian coordinates by:

$$p^C = (1 - r_1)\mathcal{T}_t^1 + \sqrt{r_1}(1 - r_2)\mathcal{T}_t^2 + \sqrt{r_1}r_2\mathcal{T}_t^3 \quad (2)$$

where, p^C is the Cartesian coordinate of point $p \in \mathcal{T}$, and $\mathcal{T}_t^1, \mathcal{T}_t^2, \mathcal{T}_t^3$ represent the *first*, *second* and *third* vertex of the t -th triangle of \mathcal{T} , respectively.

In order to perform triangulation we used the Delaunay Triangulation algorithm from the JTS Topology Suite³.

By using a triangulation of the free-of-obstacle region, or any form of parametrization, we have the benefit of avoiding collision detection checks, that is whether any sampled point is

³ver. 1.12 running in Java 1.6,
<http://tsusiatsoftware.net/jts/main.html>

inside the free region, thus making the search of bifurcation points p_1 and p_2 efficient.

An example of a triangulated free region is shown by Fig. 7, first column, where the *searchable region* \mathcal{T} is defined as the set of four Delaunay triangles between X_2 and X_3 .

Bifurcation (Fig. 7, 2nd column). In order to find the bifurcation points of the bundle $X_i = (p_1, p_2, X_j, X_k)$, we search for two points p_1 and p_2 inside the *triangulated region* \mathcal{T} by minimising the following cost function:

$$J = d(p_1, p_2) + d(p_1, X_j^1) + d(p_1, X_k^1) + d(p_2, X_j^2) + d(p_2, X_k^2), \quad (3)$$

where the subscript $d(\cdot)$ denotes the distance of the shortest path between the points encoded by the first and the second arguments; X_j^1 and X_k^1 represent the bifurcation point p_1 of the bundles X_j and X_k , respectively; and X_j^2 and X_k^2 represent the bifurcation point p_2 of the bundles X_j and X_k , respectively. Note that the distance function $d(\cdot)$ is different from Eq. 1.

For the minimization of J , we use a Matlab implementation of a convex optimization algorithm with the default parameters⁴. In order to ensure global optimality, since the presence of obstacles induces a multimodal landscape, the initial points for p_1 and p_2 are computed by the combination of the centroid coordinates of pairs of triangles drawn from the set $\mathcal{T} \times \mathcal{T}$.

An example of two bifurcation points that minimize the cost function J is shown by Fig. 7, second column. Note that these bifurcation points implicitly encode the *trunk* of the newly constructed bundle X_4 .

Fermat Points (Fig. 7, 3rd-4th column). Note that the search of the bifurcation points p_1 and p_2 for the bundle $X_i = (p_1, p_2, X_j, X_k)$ is done in a relatively small space: between the *trunks* of X_j and X_k . In order to update the structure of the bundle X_i holistically, the children for the new

⁴fmincon, with trust-region-reflective algorithm, and convergence limit on the cost function = 10^{-6}

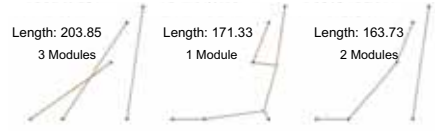


Figure 8: Example of modularity formation

trunk of X_i are computed as Fermat points in the triangles formed by p_1 and the *leaves* of X_j , and p_2 and the *leaves* of X_k . The Fermat points are inside the triangle minimise total distance to all vertices, and they are computed by the convex optimization algorithm in Matlab⁴ using initial points at the centroid coordinates of the triangles of the leaves of X_j and X_k m, respectively. By computing the Fermat points, we aim at avoiding the need to re-cluster the edges every time the topology of the tree structure \mathcal{T} changes due to bundling operations.

An example on how to update the structure of the newly formed bundle X_4 is shown by Fig. 7 (third and fourth column). Note that after finding the bifurcation points p_1 and p_2 , the bundle X_4 is still suboptimal as shown by Fig. 7 (second column). After updating the Fermat points, we can easily observe that the structure of the bundle X_4 is more compact. Also, note that while the bundling procedure is bottom-up, the update of the structure of a bundle is top-down; but both processes obey the order of the hierarchical labeling of the tree \mathcal{T} .

Finally, the sixth line of Algorithm 4 implements a *desirability criteria* a posteriori: a *null* bundle is returned in case that the resulting bundle X_i is larger than the bundle compositors X_j and X_k . Combining this simple control rule along with the generative tree representation of \mathcal{T} brings the implicit benefit to construct the modular bundles automatically. An example of this fact is shown in Fig. 8 where edge bundling without the sixth line of Algorithm 4 (Fig. 8, second column) brings a compact but longer tree structure compared to the modular and shorter bundles of Fig. 8, third column.

4. STUDY CASES

In order to evaluate the performance and behaviour of the algorithm and the generative representation, we conducted experiments consisting of different cases in the number of edges. Since we lack benchmark problems for route bundling considering obstacles and optimality criteria, we generated random instances of bipartite graphs in 2D. The future work should aim at building more sophisticated benchmark schemes.

In particularly, we used four cases considering different number of edges in randomly generated bipartite graphs:

- Case 1. Considers a bipartite graph with 10 edges whose vertices are randomly distributed over a plane.
- Case 2 and Case 3. Consider bipartite graphs with 100 edges whose vertices are randomly distributed over two sides of a plane.
- Case 4. Considers a bipartite graph with 1000 edges whose vertices are randomly distributed over two sides of a plane.

We distributed the vertices over two sides of a plane for simplicity and clarity to show the plots: in different experiments we noted that fully random placement of vertices of a graph with large number of edges results in high number of modules. We used the same obstacles in order to compare the behaviour of different graph configurations on the final bundle structure.

Fig. 9 shows the graph configuration in the left side and the main results of the bundling operation with and without including obstacles. Additional notes such as graph features, length, and the number of modules are indicated on top side of the figures. Vertices are dark-colored and each module has different color. Initial bipartite graphs are colored with the same color.

The time incurred for route bundling without obstacles for the first, second, third and

fourth case was 2.43s, 38.16s, 29.25s and 276s; while the time incurred with obstacles was 5.36s, 67.36s, 41.78s and 824.53s. Note that the time for bundling the graphs when there is a single rendered module is larger compared to the graphs with many modules. This occurs because the bundling algorithm avoids using the complete tree structure \mathcal{T} (the *desirability criteria* of the sixth line of Algorithm 4 induces not to use the entire tree \mathcal{T} in cases where complete bundling is suboptimal). Thus, bundles that render a single module will take more time compared to many modules.

In all cases, we note that it is possible to obtain compact structures with smaller length even though the existing asymmetric features of the obstacles and the positions of the vertices in the bipartite graph induce local optima issues for routing. Note that in some cases (1, 3 and 4) the structure of modularity is not preserved once obstacles are introduced, and in other cases (1 and 2) the length of the bundled routes are similar in spite of introducing obstacles in the map. Finally, besides the threshold to evaluate the convergence of the cost function J in the optimization algorithm⁴, these results are free of any other problem-specific parameter, thus being generalizable to other graph instances.

5. CONCLUSION

We have proposed a simple approach to build edge bundles using a generative grammar representation. The unique point of our approach is that not only the optimal bundles can be modeled, but also the modularity of bundles is allowed to emerge. Our main idea is to find a wiring structure that minimises the connectivity effort, expressed as the sum of the Euclidean distances among the bifurcation points, the source and target nodes in the graph. Finding such control points is relatively faster and straightforward once the generative grammar is found. Further studies remain on the agenda, including the clustering problem and the evaluation on the large scale environments.

REFERENCES

- [1] C. Cherniak, et. al, "Global optimization of cerebral cortex layout", PNAS, Vol. 101, No. 4, pp. 1081-1086, 2002.
- [2] C. Cherniak, "Component placement optimization in the brain", The Journal of Neuroscience, Vol. 14, No. 4, pp. 2418-2427, 1994.
- [3] C. Cherniak, "Neural component placement", Trends in Neurosciences, Vol. 18, No. 12, pp. 522-527, 1995.
- [4] H. Cuntz, A. Borst and I. Segev, "Optimization principles of dendritic structure", Theoretical Biology and Medical Modeling, Vol. 4, No. 21, doi:10.1186/1742-4682-4-21.
- [5] H. Cuntz, A. Mathy and M. Häuser, "A scaling law derived from optimal dendritic wiring", Prog. Brain Res., Proceedings of the National Academy of Sciences, Vol. 109, No. 27, pp. 11014-11018.
- [6] J. Clune, J.B. Mouret and H. Lipson, "The evolutionary origins of modularity", Proceedings of the Royal Society B, Vol. 280, No. 1755, pp. 1081-1086, 2002.
- [7] H. Zhou, P. Xu, X. Yuan and H. Qu, "Edge bundling in information visualization", Tsinghua Science and Technology, Vol. 18, No. 2, pp. 145-156.
- [8] D. Phan, et. al, "Flow Map Layout", Proceedings of the Proceedings of the 2005 IEEE Symposium on Information Visualization, pp. 219-224, 2005.
- [9] D. Holten and J. J. Van Wijk, "Force Directed Edge Bundling for Graph Visualization", Eurographics, Vol. 28, 2009.
- [10] E. R. Ganser, Y. Hu, S. North and C. Scheidegger, "Multilevel Agglomerative Edge Bundling for Visualizing Large Graph", IEEE Pacific Visualization Symposium, 2011.
- [11] O. Ersoy, C. Hurter, F. V. Paulovich, G. Cantareira and A. Telea, "Skeleton-Based Edge Bundling for Graph Visualization", IEEE Transactions on Visualization and Computer Graphics, Vol. 17, No. 12, pp. 2364-2373, 2011.
- [12] C. Hurter, O. Ersoy and A. Telea, "Graph Bundling by Kernel Density Estimation", Eurographics, Vol. 31, 2012.
- [13] R. Osada, T. Funkhouser, B. Chazelle, and D. Dobki, "Shape Distributions", Eurographics, ACM Transactions on Graphics, Vol. 21, No. 4, pp. 807-832, 2002.

ABOUT THE AUTHORS

1. Victor Parque is a Post-Doctoral Fellow at the Graduate School of Engineering, Toyota Technological Institute. His research interests are optimization, machine learning.
2. Masakazu Kobayashi is an associate professor at the Graduate School of Engineering, Toyota Technological Institute. His research interests are structural optimization for compliant mechanisms and design optimization for conceptual design phase.
3. Masatake Higashi is a distinguished Professor at Graduate School of Engineering, Toyota Technological Institute. His research interests are geometric modelling, product modelling, computer aided geometric design, and design and manufacturing systems.

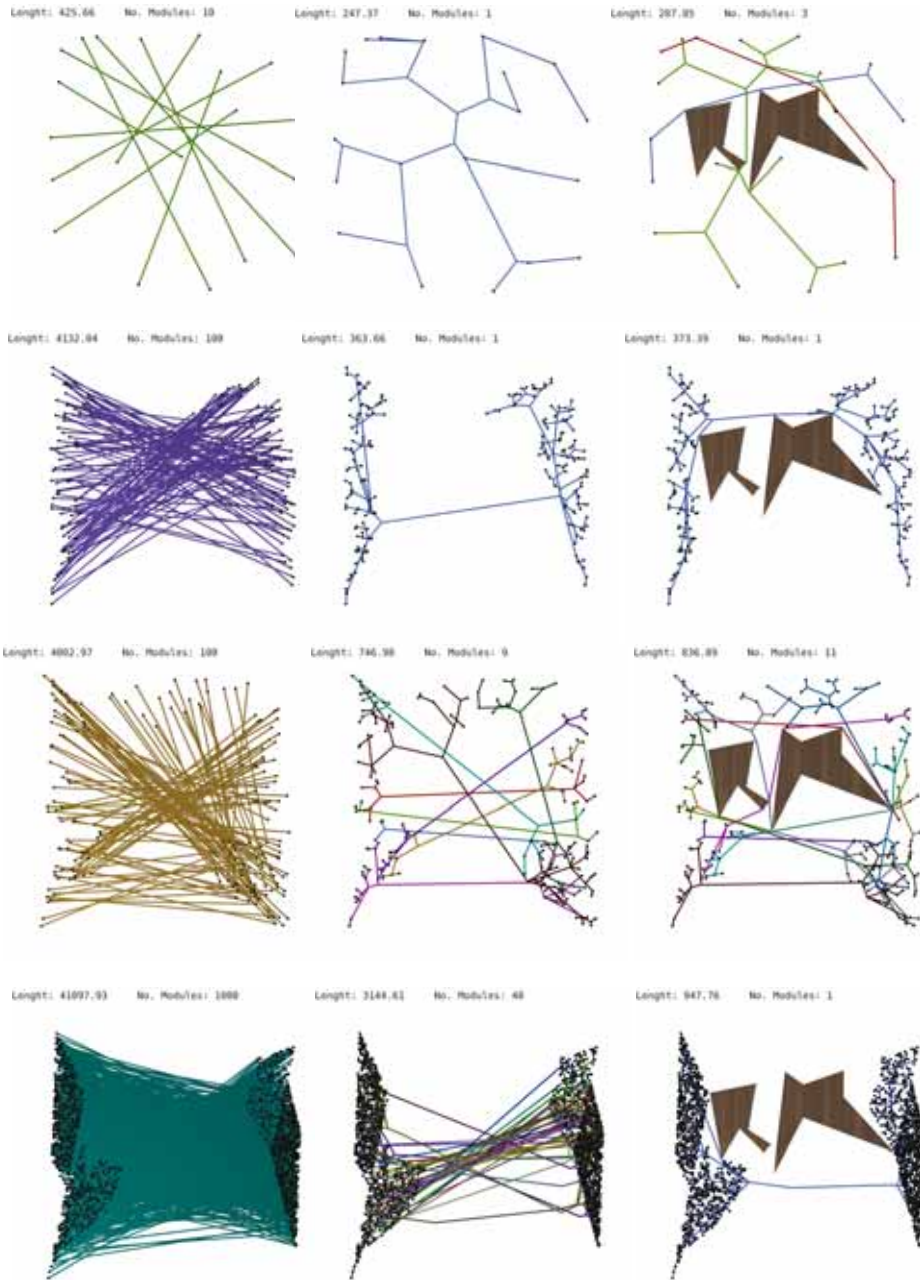


Figure 9: Four cases considering obstacle presence in varying number of edges

OVERCONSTRAINED MECHANISMS BASED ON SPECIAL PLANAR CHAINS

Otto RÖSCHEL

Graz University of Technology, Austria

ABSTRACT:

It is the aim of this paper to generalize the construction of highly overconstrained mechanisms consisting of rigid bodies linked by 1R- or spherical 2R-joints (a 2R-joint is called “spherical” if its two rotary axes intersect in a point). Based on the famous Heureka mechanism and its generalisations (see H. Stachel [13], [14] and O. Röschel [6], [7], [9]) the so-called Fulleroid mechanisms were studied in the last few years by G. Kiper et al. ([3], [4], [5]), K. Wohlhart ([15], [16], [17]) and O. Röschel ([10], [11], [12]).

The considerations are based on an observation on special planar four-bar mechanisms: In Section 1 we define special four-bar mechanisms based on quadrilaterals (quads) and work out an interesting property that can be generalized to planar polygons (Section 2). In Section 3 the planar results are embedded into space and yield further generalisations of the so-called Fulleroid linkages. According to the Chebyshev-Grübler-Kutzbach formula we compute the theoretical degree of freedom of these linkages. Our construction uses planar sub-mechanisms - therefore this theoretical value has to be adapted. In Section 4 we present some kind of modified Chebyshev-Grübler-Kutzbach formula. In Section 5 further examples of this new type of overconstrained mechanisms are presented.

Keywords: Kinematics, robotics, Fulleroid-mechanism, self-motion, generalisations of Fulleroid-mechanism, modified Chebyshev-Grübler-Kutzbach formula

1. AN OBSERVATION IN PLANAR KINEMATICS

We recall some results from paper [12] on special four-bar mechanisms based on a quadrilateral (“quad”) Q in the Euclidean plane: Let $\mathbf{a}_0, \mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ be the vertices of the quad Q . The midpoints of the pairs of points $\mathbf{a}_{i-1}\mathbf{a}_i$ (with indices modulo 4) are the points $\mathbf{b}_i := (\mathbf{a}_{i-1} + \mathbf{a}_i)/2$ – see Figure 1.¹ These midpoints $\mathbf{b}_0, \dots, \mathbf{b}_3$ form a parallelogram with sides parallel to the diagonals of Q .

The eight points \mathbf{a}_i and \mathbf{b}_i define four triangles $\mathbf{b}_i\mathbf{a}_i\mathbf{b}_{i+1}$ (rigid bodies Σ_i shaded in Figure 1) which are linked via rotary joints in the points \mathbf{b}_i .

This yields a planar parallel four-bar mechanism (see W. Wunderlich [18]) which we call *parallel (or special) four-bar mechanism FBM(Q) based on the quad Q*. The mechanism $\text{FBM}(Q)$ admits two different self-motions: One is the motion of an anti-parallelogram whilst the other one is a pure translation with circular paths. In this manuscript we will refer to that translational part only and call it $\zeta(Q, t)$. Under $\zeta(Q, t)$ the points \mathbf{b}_i form parallelograms at any moment $t \in I \subset \mathbb{R}$ of the motion.

According to [12] we describe the self-motion $\zeta(Q, t)$ as the composition of partial one-parametric motions $\zeta_i(t)$ of the bodies Σ_i with respect to a fixed *world-coordinate frame* Σ^* . We denote the rotational angles of ζ_i by ϕ_i ($i = 0, \dots, 3$). As opposite bars of a parallel four-

¹In the 3-dimensional Euclidean space (or the plane) we use coordinates with respect to a given Cartesian coordinate frame.

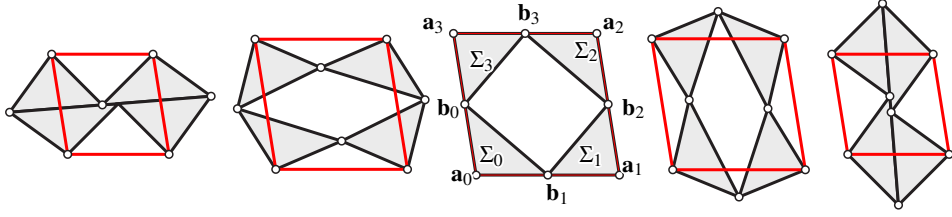


Figure 1: The quadrilateral $Q = \mathbf{a}_0, \mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ and the corresponding four-bar mechanism at several positions of the parallel self-motion.

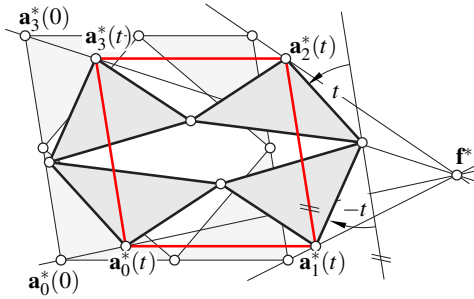


Figure 2: Two positions of $\zeta(Q, t)$ and the center of homothety.

bar mechanism rotate with the same angular velocity we can take

$$\begin{aligned} \phi_0(t) &= \phi_2(t) := t \\ \phi_1(t) &= \phi_3(t) := -t \end{aligned} \quad (1)$$

with $t \in I := [-\pi, \pi) \subset \mathbb{R}$. The rotations $\rho_i(t)$ of Σ_i about the center \mathbf{a}_i are parametrized by

$$\begin{aligned} \rho_i(t) : \Sigma_i &\rightarrow \Sigma^* \\ \mathbf{p} &\rightarrow \rho_i(\mathbf{p}, t) := \mathbf{a}_i + R(\phi_i(t))(\mathbf{p} - \mathbf{a}_i) \end{aligned} \quad (2)$$

with

$$R(\phi_i(t)) = \begin{pmatrix} \cos t & (-1)^{i+1} \sin t \\ (-1)^i \sin t & \cos t \end{pmatrix}. \quad (3)$$

They determine $\zeta_i(Q, t)$ apart from some further translational parts. Figure 2 shows two poses of the four-bar mechanism. The segments $\mathbf{a}_1\mathbf{b}_2$

and $\mathbf{b}_2\mathbf{a}_2$ are rotated by $\pm t$. At $t = 0$ we have $\mathbf{a}_i^*(0) := \rho_i(\mathbf{a}_i, 0) = \mathbf{a}_i$ ($i = 0, \dots, 3$). The straight lines through the initial points $\mathbf{a}_1^*(0), \mathbf{a}_2^*(0)$ and the corresponding positions $\mathbf{a}_1^*(t), \mathbf{a}_2^*(t)$ are parallel for all $t \in I$. Their distances are scaled by the factor

$$\rho(t) = \cos t. \quad (4)$$

As the same holds for the other pairs of coupler points and their positions the quads $Q^*(t)$ formed by the coupler points are similar to the initial one. At any moment there exists a corresponding center of homothety.

Now we add translational parts of the partial motions such that this center of homothety is kept in place. We choose a point \mathbf{f}^* in Σ^* and define the partial motions via

$$\begin{aligned} \zeta_i(t) : \Sigma_i &\rightarrow \Sigma^* \\ \mathbf{p} &\rightarrow \zeta_i(\mathbf{p}, t) := \mathbf{a}_i \cos t + \mathbf{f}^*(1 - \cos t) + \\ &+ R(\phi_i(t))(\mathbf{p} - \mathbf{a}_i). \end{aligned} \quad (5)$$

With $\mathbf{b}_i = (\mathbf{a}_{i-1} + \mathbf{a}_i)/2$ we get $\zeta_{i-1}(\mathbf{b}_i, t) - \zeta_i(\mathbf{b}_i, t) = \mathbf{o}$ for all $t \in I$ ($i = 0, \dots, 3$) (see [12]). The point paths of \mathbf{b}_i under $\zeta_{i-1}(t)$ and $\zeta_i(t)$ coincide $\forall t$. So, the partial motions $\zeta_i(t)$ (5) can be used to describe the translational self-motion $\zeta(Q, t)$ of the mechanism. These partial motions $\zeta_i(t)$ have parametrisations by linear combinations of $\{1, \sin t, \cos t\}$ - they are Cardan-motions (see O. Bottema - B. Roth [2]). This way we have generated the translational self-motion

$\zeta(Q, t)$ of this four-bar mechanism by composing four linked Cardan motions $\zeta_i(t)$. The point paths of \mathbf{a}_i under $\zeta_i(t)$ are

$$\mathbf{a}_i^*(t) = \zeta_i(\mathbf{a}_i, t) = \mathbf{a}_i \cos t + \mathbf{f}^*(1 - \cos t). \quad (6)$$

These paths are parts of straight lines through the center of homothety \mathbf{f}^* – see Figure 2. This observation yields (compare [12])

Theorem 1.1 *Given a quad $Q = \mathbf{a}_0\mathbf{a}_1\mathbf{a}_2\mathbf{a}_3$ in the Euclidean plane and the parallel four-bar mechanism $\text{FBM}(Q)$ based on this quad. Then for any choice of a point \mathbf{f}^* the described translational self-motion $\zeta(Q, t)$ ($t \in I$) of this special planar four-bar mechanism moves the four coupler points such that their positions at any time $t \in I$ form quadrilaterals $Q^*(t)$ homothetic to the initial quad. The center of these homotheties is the point \mathbf{f}^* , the scaling factor is given by (4).*

Remark 1.2 *Due to our construction we have $Q^*(-t) = Q^*(t) \forall t \in I$. The quads $Q^*(t)$ and $Q^*(-t)$ result in the same series of quads for all t .*

2. FURTHER OBSERVATIONS IN PLANAR KINEMATICS

Theorem 1.1 offers a possibility to reproduce scaled copies of an initial quad Q . Let $\mathbf{a}_0, \mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4, \mathbf{a}_5$ be the vertices of a polygon P that comprises the two quads $Q_1 := \mathbf{a}_0, \mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ and $Q_2 := \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4, \mathbf{a}_5$ – see Figure 3. Due to Chapter 1 the two quads Q_1 and Q_2 yield parallel four-bar mechanisms that can be glued together along common sides of the coupler triangles. The corresponding mechanism consists of six rigid bodies interlinked by six revolute joints. It admits a one-parametric self-motion $\zeta(P, t)$ stemming from the translational self-motions or the partial four-bars. The positions of the coupler points $\mathbf{a}_0^*(t), \dots, \mathbf{a}_5^*(t)$ define hexagons $P^*(t)$ homothetic to the initial one at any moment of the self-motion $\zeta(P, t)$.

Figure 3 gives an example. Each of the initial quads shows up in a scaled copy in each position of the mechanism's self-motion.

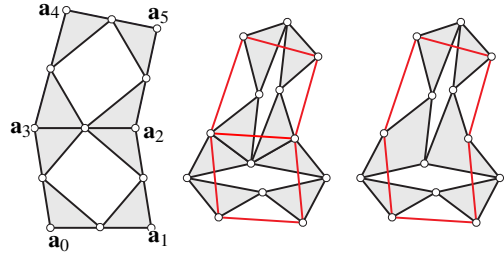


Figure 3: Two combined four-bar mechanisms.

Remark 2.1 *If we leave the combined four-bar mechanism unchanged we can watch partial sets of these coupler points (e.g. a triangle T of coupler points). Their positions under $\zeta(P, t)$ will define polygons $T^*(t)$ homothetic to the initial configuration.*

This procedure can be iterated for any point set with odd number ² $n \geq 3$. The vertices define a polygon $P := \{\mathbf{a}_0, \dots, \mathbf{a}_n\}$: This set can be arranged in quadruples Q_1, \dots, Q_k sharing pairs of vertices. We follow the procedure of Chapter 1 and get a set of k four-bar mechanisms that can be linked as the ones before. The corresponding planar mechanism will be called 'composed four-bar mechanism of the given polygon P ' $\text{CFBM}(P)$.³ It admits a translational self-motion $\zeta(P, t)$. Under $\zeta(P, t)$ the coupler points stemming from the initial polygon P define positions $P^*(t)$ which are homothetic to P for all positions of the mechanism. According to (4) the scaling factor is given by $\cos t$.

So we can sum up

Theorem 2.2 *Given a polygon $P = \mathbf{a}_0, \dots, \mathbf{a}_n$ in the Euclidean plane and one of the corresponding composed four-bar mechanisms $\text{CFBM}(P)$. Then the described translational self-motion $\zeta(P, t)$ ($t \in I$) of this special planar mechanism moves the*

²If n happens to be even, we can add an arbitrary further vertex.

³This composed four-bar mechanism related to the initial polygon P is not unique. It obviously depends on the choice of the quadruples within the set of vertices.

coupler points such that their positions at any time form polygons $P^*(t)$ homothetic to the initial one.

Remark 2.3 Figure 4 displays a configuration of three linked four-bar mechanisms. Not all possible eight coupler points are used there. Four of the corresponding coupler points form squares $P^*(t)$ at any moment of the self-motion of this mechanism.

Remark 2.4 As in Remark 1.2 we have $P^*(-t) = P^*(t) \forall t \in I$. The polygons $P^*(t)$ and $P^*(-t)$ coincide for all t .

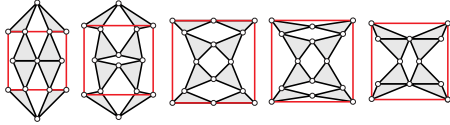


Figure 4: Some positions of three combined four-bar mechanisms – the positions of four coupler points form squares at any moment of the self-motion.

3. EXTENSION INTO SPACE

Given a polyhedron Π in three-dimensional Euclidean space \mathbb{E}_3 . We consider two of its faces with a common edge connecting two vertices $\mathbf{v}_1, \mathbf{v}_2$. Let the corresponding polygons $P_{1,2}$ have vertices $\mathbf{a}_{0,1} := \mathbf{v}_1, \mathbf{a}_{1,1} := \mathbf{v}_2, \dots, \mathbf{a}_{n_1,1}$ and $\mathbf{a}_{0,2} := \mathbf{v}_1, \mathbf{a}_{1,2} := \mathbf{v}_2, \dots, \mathbf{a}_{n_2,2}$ with $n_1, n_2 \geq 3$. The planes of these two faces are $\varepsilon_{1,2}$ – see Figure 5 for a truncated quadratic pyramid as basic polyhedron Π . For the two polygons P_i ($i = 1, 2$) in the planes ε_i (a square and a trapezoid in Figure 5) we apply the procedure presented in the last Chapters above. We generate corresponding four-bar mechanisms $FBM(P_i)$ or composed four-bar mechanisms $CFBM(P_i)$ with parallel self-motions $\zeta_i(P_i, t_i)$ in the planes ε_i . Under $\zeta_i(P_i, t_i)$ the coupler points of $CFBM(P_i)$ describe configurations $P_i^*(t_i)$ homothetic to the given polygon P_i . According to (4) the scaling factors of

$\zeta_i(P_i, t_i)$ are given by $\cos t_i$. They coincide if we put $t_2 = \pm t_1$. In these cases the polygons $P_1^*(t_1)$ and $P_2^*(\pm t_1)$ can be seen as faces of a polyhedron $\Pi^*(t_1)$ homothetic to the initial one – see Figure 5. The partial planar mechanisms $FBM(P_i)$ or $CFBM(P_i)$ can be parametrized via $t \in [-\pi, \pi)$ if we put $t_1 := t$ and $t_2 := \pm t$. The positions of the planes $\varepsilon_i^*(t)$ of the planar mechanisms contain the corresponding faces of $\Pi^*(t)$ – they keep constant angle during this one-parametric motion parametrized by t .

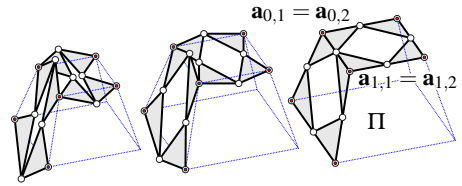


Figure 5: Two faces of a polyhedron and corresponding four-bar mechanisms.

We interlink these two motions via the common points on the common edge of the polyhedron. To do so, we can interlink the two systems by a spherical $2R$ -joint with its two axes orthogonal to the planes $\varepsilon_1, \varepsilon_2$. This additional joint will not lock the motions. The axes of the rotary joints keep constant angle and intersect in the common points $\mathbf{a}_{0,1} = \mathbf{a}_{0,2}$ and $\mathbf{a}_{1,1} = \mathbf{a}_{1,2}$, respectively.

Now we proceed with the other faces of the basic polyhedron in the same way: For each polygon P_i of Π our procedure yields planar mechanisms $FBM(P_i)$ or $CFBM(P_i)$ reproducing the shape of P_i . They can be interlinked to the neighbour mechanisms like before.

At any vertex of the polyhedron we get at least three different spherical $2R$ -joints which interlink parts stemming from the partial planar mechanisms sharing one edge of the polyhedron. The axes of these spherical $2R$ -joints share a common point (the corresponding vertex \mathbf{v}_j of the polyhedron Π). In any vertex these rotary axes form edges of a rigid pyramid that are orthogonal to the faces of Π meeting at the vertex \mathbf{v}_j .

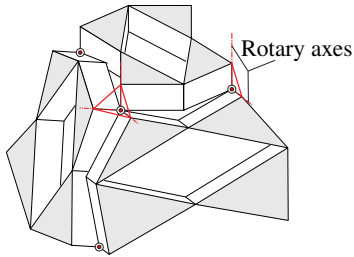


Figure 6: Three faces of the polyhedron and corresponding four-bar mechanisms. The rotary axes keep constant angles during the self-motion.

Figure 6 shows one of these pyramids for the mechanism based on the truncated pyramid from our example in red color. Our construction guarantees at least one-parametric movability of the mechanism. In total, the mechanism is built up from six parts which perform one-parametric self-motions (stemming from planar four-bar mechanisms in the six planes of Π) and of eight rigid bodies (pyramids like in Figure 6). In total, we have 24 rotary joints with axes meeting in the edges of the pyramids (the rotary axes of the four-bar mechanisms are not of interest here). This mechanism admits at least the one-parametric motion described above (parameters $t_i = \pm t$ with $t_1 := t$). According to physical restrictions this self-motion is confined to some range of $t \in [-T_1, +T_2]$ with suitable positive reals T_1, T_2 . Figure 7 shows some positions of the mechanism's self-motion.

This procedure is widely applicable and can be carried out with any polyhedron Π of the Euclidean space. In all of its faces (polygons P_i) we can determine suitable compositions $CFBM(P_i)$ of planar four-bar mechanisms that deliver coupler polygons homothetic to P_i during the corresponding self-motions. They can be linked to rigid pyramids stemming from the polyhedron vertices V_j via rotary joints. The axes of these rotary joints follow the edges of these pyramids and are orthogonal to the faces of the polyhedron Π meeting at the corresponding vertex. The

corresponding mechanism $M(\Pi)$ admits at least the one-parametric self-motion guaranteed by the construction. But it has to be checked whether there exists a physical implementation of these self-motions without self-interference. If necessary, some kind of offset can be applied - see Figure 8. In that figure the realization of all partial planar four-bars was shifted 'off' the initial polyhedron. The former pyramids of the rotary axes linking the partial motions are replaced by rigid strips (red color in Figure 8).

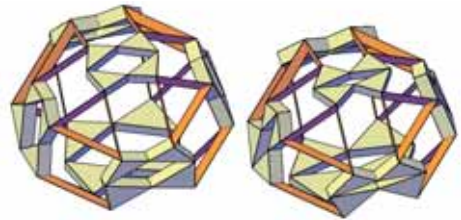


Figure 8: Two positions of the self-motion of the offset to the mechanism based on a truncated quadratic pyramid.

We can sum up:

Theorem 3.1 *Given a polyhedron Π with faces P_i and vertices \mathbf{v}_j ($i = 1, \dots, f$, $j = 1, \dots, v$). The described algorithm delivers f one-parametric planar motions $FBM(P_i)$ or $CFBM(P_i)$ in the faces of Π that can be interlinked via rotary joints to rigid pyramids corresponding to the vertices \mathbf{v}_j of Π . The emerging mechanism $M(\Pi)$ admits a one-parametric self-motion at least. In all of its poses the positions of the coupler points of the planar motions $FBM(P_i)$ or $CFBM(P_i)$ determine vertices of a polyhedron $\Pi^*(t)$ homothetic to the initial polyhedron Π .*

Remark 3.2 *These new examples can be viewed as generalisations of Fulleroid mechanisms (see K. Wohlhart [17], G. Kiper [4] and O. Röschel [10], [11], [12]).*

Remark 3.3 *According to [12] the homothetic polyhedra $\Pi^*(t)$ can be positioned with respect*

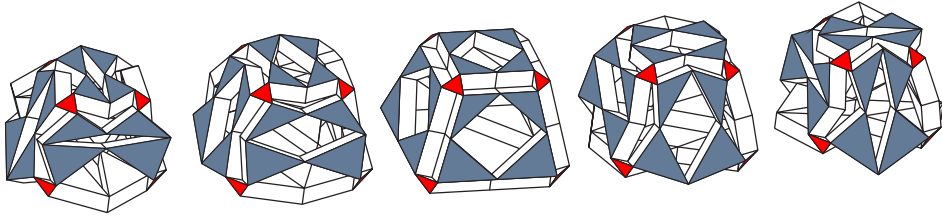


Figure 7: Some positions of the self-motion of the movable spatial mechanism based on a truncated quadratic pyramid.

to a common center \mathbf{f}^* fixed in space. Then the rigid parts of the mechanism perform Darboux-motions with respect to the corresponding fixed coordinate frame. The relative motions of different bodies are special one-parametric motions studied by the author in [8]. In general they are rational of degree 4 (compare with [12]).

The basic polygons P_i define some one-parametric planar motions $FBM(P_i)$ or $CFBM(P_i)$ parametrized by $t \in [-T_1, T_2] \subset I$. The extremal sizes of $P_i^*(t)$ are obtained at values $t \in \{-\pi, -\pi/2, 0, \pi/2\}$. According to Remark 1.2 we have bifurcations of the self-motions at these positions: For each of the partial motions $\zeta_i(P_i, t_i)$ we can reverse the direction of rotation at that moments. According to Remarks 1.2 and 2.4 this can be performed without blocking the mechanism.⁴

As one partial motion $\zeta_1(P_1, t_1 := t)$ has to be used as reference, we have 2^{f-1} different branches of the self-motion. Thus we have

Theorem 3.4 *The kinematic chain based on the polyhedron Π theoretically admits a one-parametric self-motion with at least 2^{f-1} different branches. The bifurcations arrive at the extremal positions of the underlying planar mechanisms.*

⁴This is a theoretical result. The self-motions of the physical model will be limited by several constraints. It can happen that not all of these theoretical options can really be implemented as bodies may interfere with each other or collide.

4. A MODIFIED CEBYSHEV-GRÜBLER-KUTZBACH FORMULA

The famous Cebyshev-Grübler-Kutzbach formula (see [1]) facilitates the computation of the theoretical degree of freedom of the mechanisms described in Theorem 3.1. Due to our construction the mechanism consists of parts stemming from planar mechanisms performing one-parametric self-motions. For each of the planar four-bar mechanisms seen as spatial mechanisms the theoretical degree of freedom takes on the value $F = -2$. In fact, each sub-mechanism admits a one-parametric self-motion. So our construction with planar sub-mechanisms generates redundancies in the classical Cebyshev-Grübler-Kutzbach formula. In order to remove these 'constructional redundancies' we want to give some 'modified version of the Cebyshev-Grübler-Kutzbach formula' for such cases.

Given a mechanism consisting of f sub-mechanisms $Z_i (i = 1, \dots, f)$ with one-parametric self-motions and of v further rigid bodies $B_j (j = 1, \dots, v)$. Additionally, we have m joints $J_k (k = 1, \dots, m)$ interlinking the rigid bodies B_j with parts of the partial mechanisms Z_i . The joint J_k is supposed to have $f(J_k)$ free joint parameters. With these preconditions we can determine some kind of theoretical degree of freedom for this mechanism in the following way:

Without joints the partial mechanism Z_i could be positioned with respect to 6 real position parameters and one additional real parameter of the corresponding self-motion. All these partial

mechanisms Z_i together with the rigid bodies B_j therefore could be positioned by $(6+1)f+6v$ real parameters. If we use one rigid body as reference body we could reach a positioning manifold of dimension $(6+1)f+6v-6=6(f+v-1)+f$. Each additional joint J_k will reduce this possible number by $6-f(J_k)$. So we end up with

$$\begin{aligned} F^* &:= 6(f+v-1)+f-\sum_{k=1}^m [6-f(J_k)] = \\ &= 6(f+v-m-1)+f+\sum_{k=1}^m f(J_k) \end{aligned} \quad (7)$$

as the 'modified theoretical degree of freedom' for our mechanisms. For our cases this formula takes the redundancies within the planar sub-mechanisms into account. Nevertheless it does not refer to other redundancies.

For our example of the mechanism based on the truncated pyramid we have $f=6, v=8, m=24$ and for all $f(J_k)=1$. F^* takes on the value $F^*=-36$. The Cebyshev-Grübler-Kutzbach formula would deliver the value $F=-54$. In each sub-mechanism (originating from the four-bar mechanisms) we have a difference of -3 compared to the outcome of our modified formula.

With the general procedure (see Theorem 3.1) based on a polyhedron Π with f faces, e edges and v vertices we arrive at a mechanism consisting of f partial one-parametric mechanisms and v rigid bodies which are interlinked via $2e$ rotary joints. The modified formula (7) then yields

$$\begin{aligned} F^* &= 6(f+v-2e-1)+f+2e = \\ &= 7f-10e+6v-6. \end{aligned} \quad (8)$$

For convex polyhedra Euler's polyhedra formula delivers $f-e+v=2$. Thus, for mechanisms based on convex polyhedra we have a modified degree of freedom given by

$$F^* = f-4e+6. \quad (9)$$

5. A FURTHER EXAMPLE

Here we add one further example which is based on the faces of a cube. The squares of the faces are reproduced by the coupler points of the composed planar mechanism presented in Figure 3.

If copies of this partial planar mechanism are placed in all six faces of the cube we end up with a mechanism as in Figure 9. Here we demonstrate an implementation without offsetting.

As the combinatorial structure of the composed spatial mechanism is that of the mechanism on base of the truncated pyramid the modified Cebyshev-Grübler-Kutzbach formula yields the same value $F^*=-36$ of the modified theoretical degree of freedom (the Cebyshev-Grübler-Kutzbach formula would deliver the value $F=-90$). Figure 10 displays some positions of a further branch of the self-motion for an offset to the mechanism presented in Figure 9.

6. CONCLUSIONS

In this paper a procedure is presented which is capable of generating highly overconstrained mechanisms in Euclidean space. The idea is based on results published in [12]. Any polyhedron can be used to generate such a mechanism with a one-parametric self-motion at least. This valuable tool delivers a great variety of new overconstrained mechanisms. The Cebyshev-Grübler-Kutzbach formula is a venerable and valuable tool to estimate the theoretical degree of freedom of a mechanism. In the particular case of the presented mechanisms though, it provides an estimate that is tad rough. This is why, in this paper, a modified version of the Cebyshev-Grübler-Kutzbach formula is developed. It can specifically be applied to mechanisms that are built up by sub-mechanisms as is the case in our examples.

REFERENCES

- [1] R. Beyer. Technische Raumkinematik. Springer, Berlin-Göttingen-Heidelberg 1963.
- [2] O. Bottema and B. Roth. Theoretical Kinematics. North-Holland Series in Applied Mathematics and Mechanics, Vol. 24. North-Holland Publishing Company, Amsterdam, New York, Oxford 1979.
- [3] G. Kiper, E. Söylemez and A.U. Kisisel. A

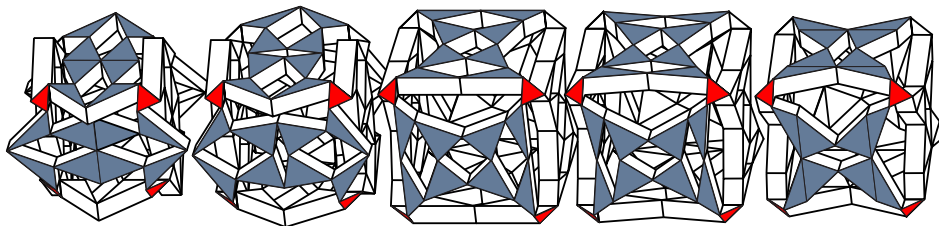


Figure 9: Some positions of the self-motion of the movable spatial mechanism based on a cube

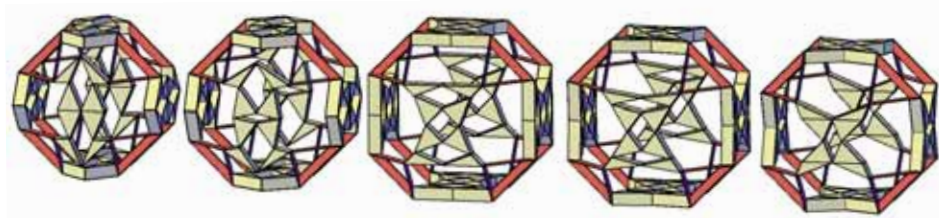


Figure 10: Some positions of another branch of the self-motion of the offset to the mechanism from Figure 9.

- family of developable Polygons and polyhedra. *Mechanisms and Machine Theory*, 43 (5), 627 - 640 (2008).
- [4] G. Kiper. Fulleroid-Like Linkages. In *Proceedings of EUROMES 08 (M. Ceccarelli ed.)*, 423-430 (2008).
- [5] G. Kiper and E. Söylemez. Homothetic Jitterbug-like linkages. *Mechanism and Machine Theory* Vol. 51, 145-158 (2012).
- [6] O. Röschel. Zwangläufig bewegliche Polyedermodelle I. *Math. Pann.* 6/1, 267 - 284 (1995).
- [7] O. Röschel. Zwangläufig bewegliche Polyedermodelle II. *Studia Sci. Math. Hung.* 32, 383 - 393 (1996).
- [8] O. Röschel. Linked DARBOUX-motions. *Math. Pannonica* 7/2, 291 - 301 (1996).
- [9] O. Röschel. Zwangläufig bewegliche Polyedermodelle III. *Math. Pann.* 12/1, 55 - 68 (2001).
- [10] O. Röschel. The self-motions of a Fulleroid-like mechanism. In *Proc. of the 14th International Conference on Geometry and Graphics*, Kyoto, August 5-9, 203, 1 - 7 (2010).
- [11] O. Röschel. Overconstrained Mechanisms Based on Trapezohedra. In *Proc. of the 15th International Conference on Geometry and Graphics*, Montreal, August 1-5, 629 - 637 (2012).
- [12] O. Röschel. Overconstrained Mechanisms Based on Planar Four-Bar-Mechanisms. *Computer Aided Geometric Design* (in print).
- [13] H. Stachel. The HEUREKA-Polyhedron. In *Proceedings of Coll. Math. Soc. J. Bolyai*, 447 - 459 (1991).

- [14] H. Stachel. Zwei bemerkenswerte bewegliche Strukturen. *Journ. of Geom.* 43, 14 - 21 (1992).
- [15] K. Wohlhart. Kinematics and Dynamics of the Fulleroid. *Multibody System Dynamics* Vol.1, 241-258 (1997).
- [16] K. Wohlhart. New Regular Polyhedral Linkages. In *Proceedings of 8th IFToMM Intern. Symposion on Theory of Machines and Mechanisms*, Bucharest, Romania, Vol II, 365-370 (2001).
- [17] K. Wohlhart. Cupola Linkages. In *Proceedings of the 12th World Congress on the Theory of Mach. and Mech.*, Besancon, France, 319-324 (2007).
- [18] W. Wunderlich. Ebene Kinematik. B.I-Hochschultaschenbücher 447/447, Bibliographisches Institut, Mannheim - Wien - Zürich 1970.

ABOUT THE AUTHOR

Otto Röschel, Dr., is Professor for geometry at the Institute of Geometry, NAWI Graz, at the University of Technology in Graz, Austria. His research interests are Kinematic Geometry, Classical Differential Geometry and Computer Aided Geometric Design. He can be contacted by e-mail: roeschel@tugraz.at or via: TU Graz, Institute of Geometry, Kopernikusgasse 24, A-8010 Graz, Austria.

THE PARABOLA IN UNIVERSAL HYPERBOLIC GEOMETRY II: CANONICAL POINTS AND THE \mathcal{Y} -CONIC

Ali ALKHALDI¹ and Norman WILDBERGER²

¹King Khalid University, Saudi Arabia

²University of New South Wales, Australia

ABSTRACT: We introduce canonical structures for the hyperbolic parabola in Universal Hyperbolic Geometry, and explicit formulas for them in Parabolic standard coordinates. Quite a few remarkable collinearities result, with the duality of the twin parabola playing a major role. We end describing the interesting \mathcal{Y} -conic which is homologous to the parabola.

Keywords: Universal hyperbolic geometry, parabola, projective geometry, canonical points, \mathcal{Y} -conic

1. INTRODUCTION AND REVIEW OF THE HYPERBOLIC PARABOLA

We begin by reviewing the basic set-up for Universal hyperbolic geometry (UHG), (see [6], [7], [8], [9]), and then the definition of a parabola in this context, and the use of *standard coordinates* which allow a significant simplification of the formulas for the parabola, following [1]. In this algebraic version of hyperbolic geometry, we use a Cayley-Klein projective framework with metrical structure determined by an invertible symmetric projective matrix C and its adjugate D . Since the theory is independent of the particular form of C , we may employ projective (linear) transformations to simplify situations. Everything holds over a general field \mathbb{F} not of characteristic two—which may for simplicity be taken to be the rational numbers. Prior classical discussions of such curves in hyperbolic geometry include [2], [3], [4], and [5].

In Figure 1 we see the null circle or absolute \mathcal{C} , the parabola \mathcal{P}_0 , and some canonical points associated to it generating the homologous \mathcal{Y} -conic.

Here is how we set this up using (projective) linear algebra. A **(projective) point** is a proportion $a = [x : y : z]$ in square brackets, or equivalently a projective row vector $a = [x \ y \ z]$. A **(projective) line** is a proportion $L = \langle l : m : n \rangle$

in pointed brackets, or equivalently a projective column vector

$$L = \begin{bmatrix} l \\ m \\ n \end{bmatrix}.$$

Incidence between the point $a = [x : y : z]$ and the line $L = \langle l : m : n \rangle$ is defined by $aL = lx + my + nz = 0$. The **join** a_1a_2 of distinct points $a_1 \equiv [x_1 : y_1 : z_1]$ and $a_2 \equiv [x_2 : y_2 : z_2]$ is the unique line passing through (i.e. incident with) a_1 and a_2 , namely

$$a_1a_2 \equiv [x_1 : y_1 : z_1] \times [x_2 : y_2 : z_2] \equiv \langle y_1z_2 - y_2z_1 : z_1x_2 - z_2x_1 : x_1y_2 - x_2y_1 \rangle. \quad (1)$$

The **meet** L_1L_2 of lines

$L_1 \equiv \langle l_1 : m_1 : n_1 \rangle$ and $L_2 \equiv \langle l_2 : m_2 : n_2 \rangle$ is the unique point lying on (i.e. incident with) L_1 and L_2 , namely

$$L_1L_2 \equiv \langle l_1 : m_1 : n_1 \rangle \times \langle l_2 : m_2 : n_2 \rangle \equiv [m_1n_2 - m_2n_1 : n_1l_2 - n_2l_1 : l_1m_2 - l_2m_1]. \quad (2)$$

Three points a_1, a_2, a_3 are **collinear** precisely when they lie on a line L ; in this case we also write $[[a_1a_2a_3]]$. Similarly three lines L_1, L_2, L_3 are **concurrent** precisely when they pass through a point a ; in this case we will also write $[[L_1L_2L_3]]$. These conditions may be directly reduced to checking that the determinant

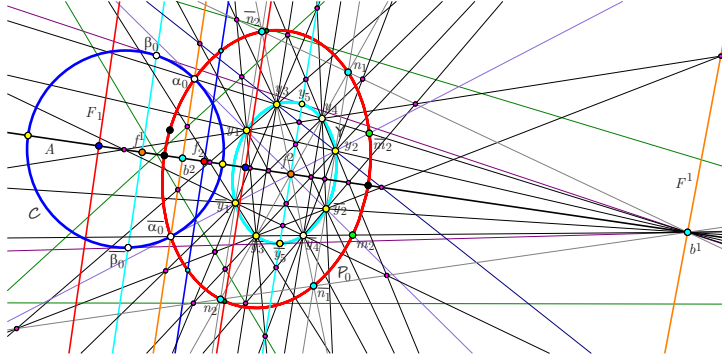


Figure 1: The hyperbolic parabola \mathcal{P}_0 and its \mathcal{V} conic

of the matrix formed by the three points or lines is zero.

1.1 Projective quadrance and spread

If C is a symmetric invertible 3×3 matrix, with entries in \mathbb{F} , and D is its adjugate matrix (the inverse, up to a multiple), denote by \mathbf{C} and \mathbf{D} the corresponding projective matrices, each defined up to a non-zero multiple. From these we get a metrical structure: the (projective) points a_1 and a_2 are **perpendicular** precisely when $a_1 \mathbf{C} a_2^T = 0$, written $a_1 \perp a_2$, and the (projective) lines L_1 and L_2 are **perpendicular** precisely when $L_1^T \mathbf{D} L_2 = 0$, written $L_1 \perp L_2$. The point a and the line L are **dual** precisely when

$$L = a^\perp \equiv \mathbf{C} a^T \text{ or } a = L^\perp \equiv L^T \mathbf{D}. \quad (3)$$

Then two points are perpendicular precisely when *one is incident with the dual of the other*, and similarly for two lines. So $a_1 \perp a_2$ precisely when $a_1^\perp \perp a_2^\perp$.

A point a is **null** precisely when it is perpendicular to itself, that is, when $a \mathbf{C} a^T = 0$, and a line L is **null** precisely when it is perpendicular to itself, that is, when $L^T \mathbf{D} L = 0$. The null points determine the **null conic**, sometimes also called the *absolute*. *Hyperbolic* and *elliptic* geometries

arise respectively from the special cases

$$C = J \equiv \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} = D \quad \text{and} \quad (4)$$

$$C = I \equiv \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = D.$$

In the hyperbolic case, $a = [x : y : z]$ is null precisely when $x^2 + y^2 - z^2 = 0$ and dually the line $L = [l : m : n]$ is null precisely when $l^2 + m^2 - n^2 = 0$. So the null circle \mathcal{C} in affine coordinates $X \equiv x/z$ and $Y \equiv y/z$ is the circle $X^2 + Y^2 = 1$, which is shown in blue in our diagrams.

In the general setting, the dual notions of **(projective) quadrance** $q(a_1, a_2)$ between points a_1 and a_2 , and **(projective) spread** $S(L_1, L_2)$ between lines L_1 and L_2 are

$$q(a_1, a_2) \equiv 1 - \frac{(a_1 \mathbf{C} a_2^T)^2}{(a_1 \mathbf{C} a_1^T)(a_2 \mathbf{C} a_2^T)} \quad \text{and} \quad (5)$$

$$S(L_1, L_2) \equiv 1 - \frac{(L_1^T \mathbf{D} L_2)^2}{(L_1^T \mathbf{D} L_1)(L_2^T \mathbf{D} L_2)}.$$

Clearly $q(a, a) = 0$ and $S(L, L) = 0$ for any point a and any line L , while $q(a_1, a_2) = 1$ precisely

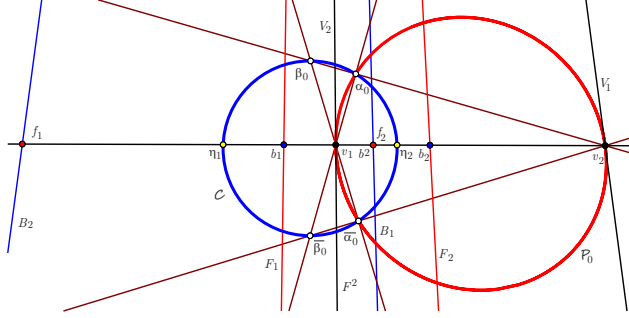


Figure 2: A parabola \mathcal{P}_0 and some basic points and lines

when $a_1 \perp a_2$, and dually $S(L_1, L_2) = 1$ precisely when $L_1 \perp L_2$. Then $S(a_1^\perp, a_2^\perp) = q(a_1, a_2)$.

In [6], Wildberger showed that for hyperbolic geometry these metrical notions agree with a purely projective formulation using suitable cross ratios, and relate to the classical hyperbolic distance $d(a_1, a_2)$ and angle $\theta(L_1, L_2)$ between points and lines, inside the null circle \mathcal{C} , via $q(a_1, a_2) = -\sinh^2(d(a_1, a_2))$ and $S(L_1, L_2) = \sin^2(\theta(L_1, L_2))$. Note however that (5) are defined for all non-null points and lines in the projective plane.

Recall also that a **midpoint** of the non-null side \overline{ab} is a point m lying on the line ab which satisfies $q(a, m) = q(m, s)$. There are generally zero or two midpoints of a given side. More novel is the following closely related concept, which we introduce in our paper [9]: a **sydpoint** of the non-null side \overline{ab} is a point s lying on the line ab which satisfies $q(a, s) = -q(b, s)$. There are also generally zero or two sydpoin of a given side, and these are intimately related to the theory of the hyperbolic parabola.

1.2 The parabola and standard coordinates

We now introduce some basic facts from [1]. The **hyperbolic parabola** \mathcal{P}_0 is defined in terms of two non-null, non-perpendicular points f_1 and f_2 (called the **foci**), as the locus of a point p_0 satisfying

$$q(p_0, f_1) + q(p_0, f_2) = 1.$$

After introducing the *directrices* $F_1 \equiv f_1^\perp$ and $F_2 \equiv f_2^\perp$ respectively, this defining equation is equivalent to either $q(p_0, f_1) = q(p_0, F_2)$ or $q(p_0, f_2) = q(p_0, F_1)$, showing that the above definition gives a hyperbolic version of the Euclidean parabola. Note also that there are *two* focus/directrix pairs.

The parabola \mathcal{P}_0 is indeed a conic. Define its **axis line** $A \equiv f_1 f_2$, the **vertices** v_1 and v_2 where \mathcal{P}_0 meets the axis, the dual **vertex lines** $V_1 \equiv v_1^\perp$ and $V_2 \equiv v_2^\perp$ which are tangents to the parabola at the vertices, and the **base points** $b_1 \equiv F_1 A$ and $b_2 \equiv F_2 A$, with dual **base lines** $B_1 \equiv b_1^\perp$ and $B_2 \equiv b_2^\perp$.

For a point $c = [x : y : z]$ its reflection in the axis A , called the **opposite** of c , is $\bar{c} = [x : -y : z]$. This is a fundamental symmetry for the parabola.

The main idea to study the parabola is to allow flexibility in our field and to *carefully choose an optimum coordinate framework*; for this we utilize four important points associated to the parabola: a pair of opposite null points $\alpha_0, \bar{\alpha}_0$ lying on \mathcal{P}_0 , and the vertices v_1, v_2 . The existence of $\alpha_0, \bar{\alpha}_0$ may well require a quadratic field extension, which we assume we have made.

We may now invoke the Fundamental theorem of projective geometry to projectively transform these four points to

$$\begin{array}{ll} \alpha_0 = [1 : 1 : 1] & \bar{\alpha}_0 = [1 : -1 : 1] \\ v_1 = [0 : 0 : 1] & v_2 = [1 : 0 : 0]. \end{array}$$

This choice is called **standard coordinates** for the parabola. It is then a pleasant fact that $\beta_0 \equiv (v_2\alpha_0)(v_1\bar{\alpha}_0)$ and $\bar{\beta}_0 \equiv (v_2\bar{\alpha}_0)(v_1\alpha_0)$ are the null points

$$\beta_0 = [-1 : 1 : 1] \quad \text{and} \quad \bar{\beta}_0 = [-1 : -1 : 1].$$

The main *Parabola standard coordinates theorem* then shows that the original hyperbolic bilinear form of $C = D = J$ from (4) is transformed to one with new matrices

$$C = \begin{pmatrix} \alpha^2 & 0 & 0 \\ 0 & 1 - \alpha^2 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

and

$$D = \begin{pmatrix} \alpha^2 - 1 & 0 & 0 \\ 0 & -\alpha^2 & 0 \\ 0 & 0 & \alpha^2(1 - \alpha^2) \end{pmatrix} \quad (6)$$

for some parameter α . While the metrical structure has now changed, the quadrance and spread depend only on the corresponding projective matrices **C** and **D**, so all the definitions of the previous section apply. Crucially, in standard coordinates the parabola \mathcal{P}_0 now has equation

$$y^2 = xz \quad (7)$$

and can be parametrized by $p_0 = p(t) \equiv [t^2 : t : 1]$. The equation of the axis in standard coordinates is $A = \langle 0 : 0 : 1 \rangle$, while the null circle \mathcal{C} is

$$\alpha^2 x^2 + (1 - \alpha^2) y^2 - z^2 = 0.$$

Almost all subsequent formulas for points, lines and related curves will involve the parameter α .

1.3 Dual conics and the connection with sydpoints

Here are the coordinates of the points and lines already defined in standard coordinates:

$$\begin{aligned} f_1 &= [\alpha + 1 : 0 : \alpha(\alpha - 1)] \\ f_2 &= [1 - \alpha : 0 : \alpha(\alpha + 1)] \\ F_1 &\equiv f_1^\perp = \langle \alpha(\alpha + 1) : 0 : 1 - \alpha \rangle \\ F_2 &\equiv f_2^\perp = \langle \alpha(\alpha - 1) : 0 : 1 + \alpha \rangle \\ b_1 &\equiv F_1 A = [\alpha - 1 : 0 : \alpha(\alpha + 1)] \\ b_2 &\equiv F_2 A = [\alpha + 1 : 0 : \alpha(1 - \alpha)] \\ B_1 &\equiv b_1^\perp = \langle -\alpha(\alpha - 1) : 0 : \alpha + 1 \rangle \\ B_2 &\equiv b_2^\perp = \langle \alpha(\alpha + 1) : 0 : \alpha - 1 \rangle. \end{aligned}$$

We define the **axis null points** to be the meets of the axis A and the null conic \mathcal{C} :

$$\eta_1 = [-1 : 0 : \alpha], \quad \eta_2 = [1 : 0 : \alpha];$$

note that this is a notational switch from [1]. We also have dual lines

$$\begin{aligned} \alpha_0^\perp &= \mathbf{C}[1 : 1 : 1]^T = \langle \alpha^2 : 1 - \alpha^2 : -1 \rangle \\ \bar{\alpha}_0^\perp &= \mathbf{C}[1 : -1 : 1]^T = \langle \alpha^2 : \alpha^2 - 1 : -1 \rangle. \end{aligned}$$

We introduce the points d_0 and \bar{d}_0 to be the meets of the directrix F_2 with the parabola \mathcal{P}_0 , should they exist, and the corresponding null points δ_0 and $\bar{\delta}_0$ lying on the directrix F_1 .

The **tangent line** to \mathcal{P}_0 at $p_0 = p(t) \equiv [t^2 : t : 1]$ on it is $P^0 = \langle 1 : -2t : t^2 \rangle$, and the dual point of this tangent line is the **twin point** p^0 of p_0 . The locus of p^0 as p_0 varies along \mathcal{P}_0 turns out, remarkably, to be another parabola \mathcal{P}^0 with foci which are the *sydpoints* f^1, f^2 of the side $\bar{f}_1 \bar{f}_2$, as in Figure 3.

To understand this, we first introduce, as in Figure 3, the lines and points

$$\begin{aligned} F^2 &\equiv \alpha_0 \bar{\alpha}_0 = \langle 1 : 0 : -1 \rangle \\ B^1 &\equiv \beta_0 \bar{\beta}_0 = \langle 1 : 0 : 1 \rangle \\ b^2 &\equiv F^2 A = [1 : 0 : 1] \\ f^1 &\equiv B^1 A = [-1 : 0 : 1]. \end{aligned}$$

The duals are

$$\begin{aligned} f^2 &\equiv (F^2)^\perp = [1 : 0 : \alpha^2] \\ b^1 &\equiv (B^1)^\perp = [1 : 0 : -\alpha^2] \\ B^2 &\equiv (b^2)^\perp = \langle -\alpha^2 : 0 : 1 \rangle \\ F^1 &\equiv (f^1)^\perp = \langle \alpha^2 : 0 : 1 \rangle. \end{aligned}$$

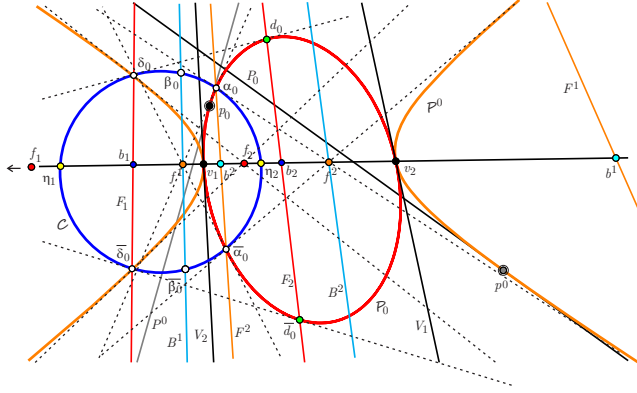


Figure 3: Twin points p_0 and p^0 on \mathcal{P}_0 and \mathcal{P}^0

The points f^1 and f^2 are the **t-foci** of the parabola \mathcal{P}_0 , while the respective dual lines F^1 and F^2 are the **t-directrices** of \mathcal{P}_0 . The meets of the t-directrices and the axis A are the **t-base points** $b^1 \equiv F^1A$ and $b^2 \equiv F^2A$, with respective dual lines B^1 and B^2 .

The *Parabola sydpoints theorem* then asserts that the points f^1 and f^2 are in fact the *sydpoints of the original side* $\overline{f_1 f_2}$. The parabola \mathcal{P}^0 with foci f^1 and f^2 , called the **twin parabola** of \mathcal{P}_0 , is the dual conic of \mathcal{P}_0 with respect to the null circle \mathcal{C} ; namely the locus of p^0 as p_0 varies. The equation of \mathcal{P}^0 in standard coordinates is

$$y^2 = \frac{-4\alpha^2}{(\alpha^2 - 1)^2} xz.$$

Note in Figure 3 that the tangents to both the parabola \mathcal{P}_0 and the null circle \mathcal{C} at their common meets, namely the null points α_0 and $\overline{\alpha_0}$, pass through the foci of the twin parabola \mathcal{P}^0 . Dually the tangents to both the parabola \mathcal{P}^0 and the null circle \mathcal{C} at their common meets, namely the null points δ_0 and $\overline{\delta_0}$ on F_1 , pass through the foci of \mathcal{P}_0 . In both cases, these tangents to \mathcal{C} are tangents to the respective twin parabola.

2. CANONICAL STRUCTURES ON THE HYPERBOLIC PARABOLA

In the paper [1] we mostly concentrated on properties of the hyperbolic parabola that were analogous to the classical theory for a Euclidean parabola. We now derive some interesting results that have no classical parallel: while a Euclidean parabola has relatively few canonical points and lines associated to it, the situation is dramatically different here, due to the existence of null points α_0 and $\overline{\alpha_0}$. Here we sketch the beginnings of this theory, up to the discussion of the *Y-conic* of a hyperbolic parabola. As usual, obtaining explicit formulae in standard coordinates is a main aim.

The proofs of most of the results are straightforward—we compute joins and meets, or duals, and occasionally verify that a point lies on the parabola \mathcal{P}_0 using (7). Many of these facts have *generalizations that are also valid when α_0 is replaced by a general point p_0 on the parabola*; we will discuss this in a future paper.

Define the **e-points**:

$$\begin{aligned} e_0 &\equiv (\eta_1 \alpha_0) (\eta_2 \overline{\alpha_0}) = [1 : \alpha : \alpha^2] \\ \overline{e_0} &\equiv (\eta_1 \overline{\alpha_0}) (\eta_2 \alpha_0) = [1 : -\alpha : \alpha^2] \end{aligned}$$

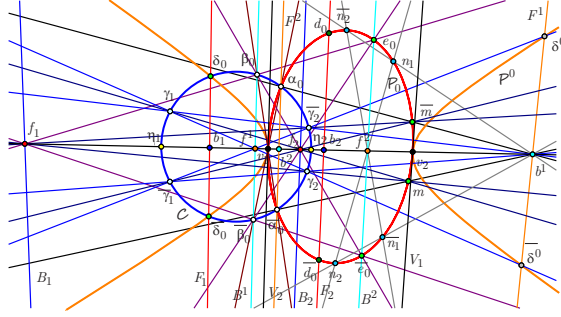


Figure 4: Additional collinearities with canonical points

the m -points:

$$m \equiv (f^2 \alpha_0) (b^1 \overline{\alpha_0}) = [1 : -\alpha^2 : \alpha^4]$$

$$\overline{m} \equiv (f^2 \overline{\alpha_0}) (b^1 \alpha_0) = [1 : \alpha^2 : \alpha^4]$$

and the n -points:

$$n_1 \equiv (f_1 \alpha_0) (b_2 \overline{\alpha_0}) =$$

$$[(\alpha + 1)^2 : -\alpha (\alpha^2 - 1) : \alpha^2 (\alpha - 1)^2]$$

$$n_2 \equiv (f_2 \alpha_0) (b_1 \overline{\alpha_0}) =$$

$$[(\alpha - 1)^2 : \alpha (\alpha^2 - 1) : \alpha^2 (\alpha + 1)^2]$$

$$\overline{n_1} \equiv (f_1 \overline{\alpha_0}) (b_2 \alpha_0) =$$

$$[(\alpha + 1)^2 : \alpha (\alpha^2 - 1) : \alpha^2 (\alpha - 1)^2]$$

$$\overline{n_2} \equiv (f_2 \overline{\alpha_0}) (b_1 \alpha_0) =$$

$$[(\alpha - 1)^2 : -\alpha (\alpha^2 - 1) : \alpha^2 (\alpha + 1)^2].$$

Theorem 1 (Canonical parabola points) *The points $e_0, \overline{e_0}, m, \overline{m}, n_1, n_2, \overline{n_1}$ and $\overline{n_2}$ all lie on the parabola \mathcal{P}_0 .*

Proof. This can be checked easily from the above forms of these points and the equation (7) for \mathcal{P}_0 . ■

These **canonical points** for the parabola \mathcal{P}_0 appear in Figure 4. Since $\alpha_0^\perp = \alpha_0 f^2 = \alpha_0 m$ and $(\overline{\alpha_0})^\perp = \overline{\alpha_0} f^2 = \overline{\alpha_0} \overline{m}$, the points m and \overline{m} are also characterized by being the respective meets of these duals with \mathcal{P}_0 .

Define the γ -points:

$$\gamma_1 \equiv (f_1 \alpha_0) (b_1 \overline{\alpha_0}) =$$

$$[\alpha^3 - \alpha^2 + \alpha + 1 : -2\alpha^2 : \alpha (\alpha^3 - \alpha^2 - \alpha - 1)]$$

$$\overline{\gamma_1} \equiv (f_1 \overline{\alpha_0}) (b_1 \alpha_0) =$$

$$[\alpha^3 - \alpha^2 + \alpha + 1 : 2\alpha^2 : \alpha (\alpha^3 - \alpha^2 - \alpha - 1)]$$

$$\gamma_2 \equiv (f_2 \alpha_0) (b_2 \overline{\alpha_0}) =$$

$$[\alpha^3 + \alpha^2 + \alpha - 1 : 2\alpha^2 : -\alpha (\alpha^3 + \alpha^2 - \alpha + 1)]$$

$$\overline{\gamma_2} \equiv (f_2 \overline{\alpha_0}) (b_2 \alpha_0) =$$

$$[\alpha^3 + \alpha^2 + \alpha - 1 : -2\alpha^2 : -\alpha (\alpha^3 + \alpha^2 - \alpha + 1)].$$

Theorem 2 (Canonical null γ -points) *The γ -points are all null points.*

Proof. We can check that each of $\gamma_1, \overline{\gamma_1}, \gamma_2$ and $\overline{\gamma_2}$ satisfy the equation $\alpha^2 x^2 + (1 - \alpha^2) y^2 - z^2 = 0$ of the null circle \mathcal{C} in standard coordinates. ■

Define the δ -points

$$\delta_1 \equiv (\beta_0 f_1) (\overline{\beta_0} b_1) =$$

$$[\alpha^3 + \alpha^2 + \alpha - 1 : 2\alpha^2 : \alpha (\alpha^3 + \alpha^2 - \alpha + 1)]$$

$$\overline{\delta_1} \equiv (\overline{\beta_0} f_1) (\beta_0 b_1) =$$

$$[\alpha^3 + \alpha^2 + \alpha - 1 : -2\alpha^2 : \alpha (\alpha^3 + \alpha^2 - \alpha + 1)]$$

$$\delta_2 \equiv (\beta_0 f_2) (\overline{\beta_0} b_2) =$$

$$[\alpha^3 - \alpha^2 + \alpha + 1 : -2\alpha^2 : \alpha (-\alpha^3 + \alpha^2 + \alpha + 1)]$$

$$\overline{\delta_2} \equiv (\overline{\beta_0} f_2) (\beta_0 b_2) =$$

$$[\alpha^3 - \alpha^2 + \alpha + 1 : 2\alpha^2 : \alpha (-\alpha^3 + \alpha^2 + \alpha + 1)].$$

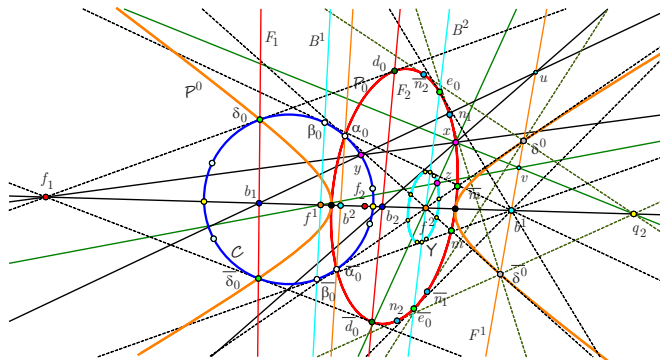


Figure 7: Homologies between \mathcal{P} and $\mathcal{P}_0, \mathcal{C}$.

Theorem 12 (\mathcal{P} conic homology) *If the meet of the tangents to \mathcal{P}_0 at n_2 and \bar{n}_2 is denoted q_2 , then the homology ϕ with axis F^1 and center f^2 which sends q_2 to f^1 sends \mathcal{P}_0 to \mathcal{P} .*

Proof. We first compute that

$$q_2 = [(\alpha + 1)^2 : 0 : -\alpha^2(\alpha - 1)^2],$$

and then determine that

$$\begin{aligned} \phi([x : y : z]) = \\ [z + 2z\alpha + x\alpha^2 - 2x\alpha^3 - x\alpha^4 - z\alpha^2 : -4y\alpha^3 : \\ \alpha^2(z - 2z\alpha + x\alpha^2 + 2x\alpha^3 - x\alpha^4 - z\alpha^2)] \end{aligned}$$

After substitution, we find that this point lies on the \mathcal{P} -conic above precisely when $64\alpha^8(xz - y^2)(\alpha^2 - 1)^2 = 0$. So this homology sends \mathcal{P}_0 to \mathcal{P} . ■

3. CONCLUSION

The theory of a hyperbolic geometry in UHG is a rich topic with many more discoveries to be made.

REFERENCES

- [1] A. Alkhaldi and N. J. Wildberger, The Parabola in Universal Hyperbolic Geometry I, *KoG*, **17**, (2013), 14-42.
- [2] G. Csima and J. Szirmai, Isoptic curves of the conic sections in the hyperbolic and elliptic plane, *Stud. Univ. Žilina, Math.Ser.* **24**, **1**, (2010), 15-22.
- [3] M. Henle, Will the Real Non-Euclidean Parabola Please Stand up?, *Math. Magazine*, **71**, No. 5 (Dec., 1998), 369-376.
- [4] E. Molnár, A hiperbolikus geometria kúpszeleteinek aszimptota tulajdonságai, az aszimptoták és fókuszok kapcsolata szintetikus tárgyalásban, *Mat.Lapok*, **22**, (1971), 77-91.
- [5] E. Story, On Non-Euclidean Properties of Conics, *American Journal of Math.*, (1883), 358-381.
- [6] N. J. Wildberger, Universal Hyperbolic Geometry I: Trigonometry, *Geometriae Dedicata*, (2013), 215-274.
- [7] N. J. Wildberger, Universal Hyperbolic Geometry II: A pictorial overview, *KoG*, **14**, (2010), 3-24.
- [8] N. J. Wildberger, Universal Hyperbolic Geometry III: First steps in projective triangle geometry, *KoG*, **15**, (2011), 25-49.
- [9] N. J. Wildberger and A. Alkhaldi, Universal Hyperbolic Geometry IV: Sydpoints and Twin Circumcircles, *KoG*, **16**, (2012), 43-62.

4. ABOUT THE AUTHORS

1. Ali Alkhalidi is a recent graduate from University of New South Wales and is now teaching at King Khalid University in Abha, Saudi Arabia.
2. Norman Wildberger is the developer of Rational Trigonometry and is an ardent YouTube poster (more than 500 maths videos at user: njwildberger).

PARAMETRIC APPROACH TO DATA INTERPOLATING CURVES USING THE SUM OF BELL-SHAPED FUNCTIONS

Iryna TARAS

Ivano-Frankivsk National Technical University of Oil and Gas, Ukraine

ABSTRACT: Transformation of discrete function into a continuous one is an important problem. The different physical nature of discrete data demands different interpolation patterns. This article gives a calculus for the solution of that interpolation problem by using the sum of bell-shaped basis functions. Continuity of the differential performances is typically to mechanical nature curves. The problem is to obtain continuity of the differential performances in the absence of point differential data. The interpolation advantages are continuity of the curve differential performances (without the additional information). Interpolating function is represented not only in an explicit form, but also in parametric one. For example, independent parameter is the length of a drilled well at well drilling check measurement. The length of a drilled well corresponds to the length of a curve at geometrical model. It has been offered to use an approximated value of the length of curve as a parameter also. The natural parameterization is convenient for the subsequent differential performances definition and also for the subsequent researches of the curve mechanical stresses. This representation is compared with representation of using the sum of different functions with bell-shaped graphs. Control parameters, which influence an oscillation of a curve, have been investigated. Test calculation for definition of the offered algorithm correctness has been made. Comparison of the curve differential performances has been fulfilled as well. Interpolations of plane and space curves have been depicted in papers.

Keywords: data-interpolation, continuous interpolation, smooth curves, parametric curve.

1. INTRODUCTION

Experimental data (or another data) are represented by discrete set of points. Such cases are frequent in engineering problem. But continuous one representation is the requisite condition for problem solution. The continuous model of such data should have geometrical properties, (differential also) which correspond to a physical source of the discrete data.

The different physical sources of discrete data demand different interpolation patterns. For instance, length of a well is the independent parameter on well drilling check measurements. The length of a well corresponds to the length of the curve at geometrical model.

Many engineering problems demand construction of the continuous curves of the de-

fined level of geometric continuity, though additional information on derivatives or curvature of the set points is absent. Using joint interpolation demands coordination or definition of the values of the derivatives in the points [1].

Constitutive advantage of the continuous interpolation is the guarantee of the order of smoothness (without the additional information). But disadvantage of this method is the length of the interpolation function. It is opportune to use the functions having limited effective interval as function-items for interpolation function. The functions with bell-shaped graphs are the following: Gauss function, function of a hyperbolic secant etc. Application of Gauss and hyperbolic secant functions for the plane and space curve interpolation has been considered in [2-4]. Usage of fractional rational

function for the curve and surface interpolation has been considered in [5-7].

2. BASIC INTERPOLATION FUNCTION

Discrete function of a curve is presented by the ordered array of knot points $\{M_i(x_i, y_i, z_i)\}$, for plane curve - $\{M_i(x_i, y_i)\}$.

Fractional rational function in the form

$$y = 1/(1 + a \cdot x^2) \quad (1)$$

is the simplest premises functions. There is only one manipulated variable a in this function form.

Using interpolating function (1) for plane curve is set down in explicit form

$$y(x) = \sum_{i=1}^n \frac{a_i}{1 + k_i \cdot (x - t_i)^2} \quad (2)$$

where: n is number of knots; k_i is manipulated variables (selected depending on the data and the interpolation pattern); t_i is a parameter of displacement of axes of function-items (their effect on result curve is considered below), a_i is coefficient of equation. They are calculated from the uniquely determined set of linear equations. That equation set is based on condition - interpolation curve is passed through all points of the array.

It is obvious the number of the set equations will increase and will enable also to compute coefficients k_i if the differential characteristics are known in knots. But that equation set will not be linear. Effective interval of the fractional rational function and Gauss one have been analyzed in [5].

Even straight line can be presented as sum of function-items (1) with displacements. Plane case is illustrated in Figure 1 (the highest line is the real straight line).

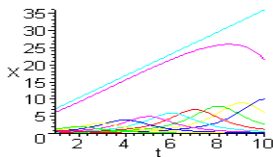


Figure 1: Example of function-items and sum interpolation function.

It is possible to use several variant of function-item displacement if function y is presented by x . It is possible to use several variants of function parameterization too.

3. INTERPOLATION FUNCTION IN CARTESIAN COORDINATE SYSTEM

The interpolation function is defined by

$$\begin{cases} x(t) = \sum_{i=1}^n \frac{a_i}{1 + k_i \cdot (t - t_i)^2} \\ y(t) = \sum_{i=1}^n \frac{b_i}{1 + k_i \cdot (t - t_i)^2} \\ z(t) = \sum_{i=1}^n \frac{c_i}{1 + k_i \cdot (t - t_i)^2} \end{cases} \quad (3)$$

for general parametric form in Cartesian coordinates. In plane case the third equation from (3) loses meaning.

It should be noted that arrays of knot points are just ordered for curves as opposed to surfaces. This fact complicates curve interpolation and influences the way to its parametrization.

Parameter t_i can possess the knot index number value 1, 2, ... n or value of x -coordinate of knot x_i . Under the condition that parameter t is equal to x interpolation function becomes explicit one. In this case the first equation from (3) loses meaning. This parameterization type is equivalent to separate modeling of two curve projections.

One more pattern of interpolation is interpolation function with the length of the curve as the independent parameter t . In this type of function parameter t_i is equal length of the curve from first knot till n_i^{th} .

3.1 Using explicit function and index number parametrization for interpolation function

From the previous researches it is known that the decreasing parameter k results in smoothing of the curve for interpolation function. It is valid for all interpolation function forms but different patterns depend in different ways on parameter k [5-7]. Interpolation result depends on data point uniformity too.

Graphic example is a straight line $y(x) = x$. There are 10 selected non-uniform knots on it as its discrete representation. Curve interpolation examples are based on small number of data points only for better visualization.

Curvature of line has been calculated for different values of parameter t_i . Results for $t_i = n_i$ and $t_i = x_i$ of similar value of parameter k_i are illustrated with graphs (fig. 2). Segment $1 \leq n \leq 10$ is the equivalent of segment $1 \leq x \leq 16$. This segment has been considered.

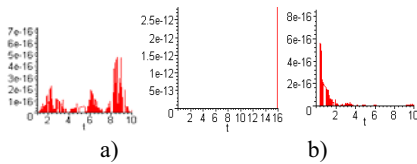


Figure 2: Curvature of interpolation function ($\kappa = 0.1$): a) $t_i = n_i$, b) $t_i = x_i$.

Acceptable results have been obtained even for non-uniform set of conical helical line points ($x = 10\varphi \cdot \cos\varphi$, $y = 10\varphi \cdot \sin\varphi$, $z = 4\varphi$) for $t_i = n_i$.

It should be beared in mind that point sequence is very important for curve interpolation. Result of the same point set interpolation using any base function (Gauss one, for example) is presented in Figure 3. This interpolation pattern is inadequate as is obvious from function graph.

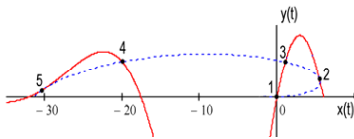


Figure 3: Plan view of the test (blue) and interpolation (red) curves using Gauss function ($\kappa = 0.01$ and $t_i = x_i$).

It is possible to transform this interpolation pattern into a suitable one for usage. Figure 4 illustrates interpolation result of the same problem situation with the only difference that the independent variable is z .

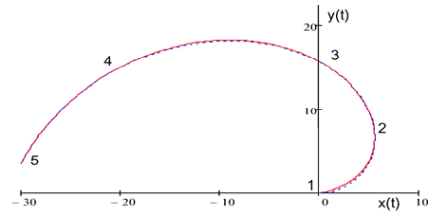


Figure 4: Plan view of the test (blue) and interpolation (red) curves using Gauss function ($\kappa = 0.01$ and $t_i = z_i$)

As statement of a problem is ordered array of points $\{M_i(x_i, y_i, z_i)\}$, it is possible to select suitable independent variable. It is necessary to use additional parameter (n , for example) if there is not monotonically increasing or decreasing sequence of coordinates.

The helical curve (radius - 2, helix lead - 8π) has been considered as an example of space curve interpolation. There are uniformly selected knots in the segments $0 \leq \varphi \leq 2$ and there are non-uniformly selected knots in $0 \leq \varphi \leq 3$. Uniform data points constitute the base for better interpolation curve then non-uniform one. It is an expected result. Curvature for uniform case is presented in Figure 5.

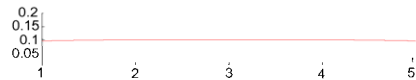


Figure 5: Curvature of interpolation function ($\kappa = 0.01$, $t_i = n_i$).

In non-uniform case curvature of interpolation lines for $t_i = n_i$ and $t_i = x_i$ can be compared from Figure 6 and Figure 7.

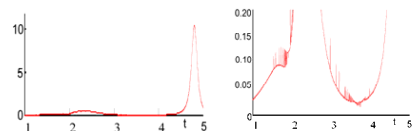


Figure 6: Curvature of interpolation function ($\kappa = 0.01$, $t_i = n_i$).

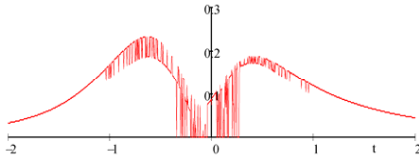


Figure 7: Curvature of interpolation function ($\kappa=0.01$, $t_i = x_i$).

Other space curve has been considered by using catenary curve as a plan view. Points were placed on this curve non-uniformly. Examined segment was $1 \leq n \leq 10$ ($-5.5 \leq x \leq 8$). View of the curves using different patterns of interpolation are illustrated in Figure 8 a) and 9 a). Curvature of these interpolation curves can be compared from Figures 8 b) and 9 b).

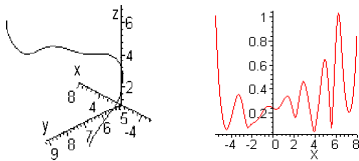


Figure 8: Interpolation curve and its curvature ($\kappa = 0.1$ and $t_i = x_i$).

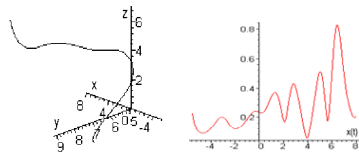


Figure 9: Interpolation curve and its curvature ($\kappa = 0.1$ and $t_i = n_i$).

The explicit and parametrical forms of the interpolation curve give equivalent results at the uniform representation of the space curve. The calculations represent this fact. The same results have been given at weak non-uniformity of the representation curve. The preference of the interpolation function form has been determined by a concrete point set and a specific problem. The better result of interpolation at $t_i = x_i$ can be received for the strong

non-uniformity of the curve representation. But the parametric form of interpolation function is a successful one at multiple-valued function (having loop on the graph etc.).

3.2 Natural parameterization of interpolation function

The problem of real length l_i is indefinable. It is possible to use only its approximate value. Polygonal line with the same knots may be used to estimate value length l_i , for instance [6].

Arc approximation can be used for the higher accuracy. It can be arc of circle passing through three points.

The previous set of points has been considered for the purpose of different interpolation pattern comparison. Figure 10 illustrates interpolation curve and its curvature at $t = l$. Length of segment l_i has been calculated approximately. Segment of curve was taken as the polygonal line segment.

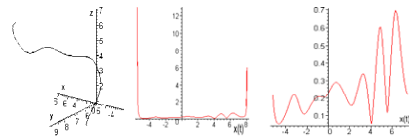


Figure 10: Interpolation curve and its curvature ($\kappa = 0.1$ and $t_i = l_i$).

Results of this interpolation pattern can be compared with Figure 8 and 9.

As follows from Figures 8, 9, 10 the interpolation pattern is the most incorrect in end points at $t_i = l_i$.

The same result has been obtained for conical helical line. This line was given using non-uniform point set described above (spiral - in plane view, lead of helix - 8π).

Figure 11 illustrates interpolation curve error in neighborhood of first knot $x = 0$ (segment at $x < 0$ is extrapolated). It should be noted that such an effect has not been observed while using explicit function and index number parameterization for the same data points.

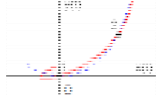


Figure 11: Plan view of the interpolation function (red) and the test curve (blue) graph in neighborhood of point $x = 0$ ($\kappa = 0.01$, $t_i = l_i$).

This pattern of interpolation is the most unequal especially for end points.

4. POLAR AND CYLINDRICAL COORDINATE SYSTEM

Polar coordinate system is unappreciated in interpolation problems. Frequently the experimental data represent a 2-D discrete set of points as array of knots $\{M_i(\rho_i, \varphi_i)\}$ in polar coordinates or a 3-D discrete set of points as $\{M_i(\rho_i, \varphi_i, z_i)\}$ in cylindrical coordinate system.

In this case the interpolation function (3) is transformed into a general form

$$\begin{cases} \varphi(t) = \sum_{i=1}^n \frac{a_i}{1 + k_i \cdot (t - t_i)^2} \\ \rho(t) = \sum_{i=1}^n \frac{b_i}{1 + k_i \cdot (t - t_i)^2} \\ z(t) = \sum_{i=1}^n \frac{c_i}{1 + k_i \cdot (t - t_i)^2} \end{cases} \quad (4)$$

Interpolation patterns have been analyzed for polar and cylindrical coordinate system. The parameter effect has been also analyzed by using the foregoing interpolation function. The decreasing parameter k results in smoothing of the curve for these kinds of coordinate system too.

Parameter $t = \varphi$ was used for the following patterns and test examples.

The same data point set of spiral $\rho = 10\varphi$ was considered by using parameter $\kappa = 0.01$. Absolute error of function and curvature of the curve are presented in Figure 12 and 13.

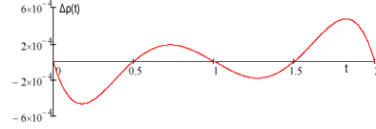


Figure 12: Error of interpolation function ($\kappa = 0.01$, $t = \varphi$).

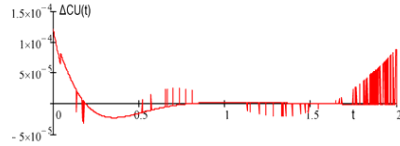


Figure 13: Error of interpolation function curvature ($\kappa = 0.01$, $t = \varphi$).

It is a predictable result for this kind of curves.

The extraordinary example is represented below. Data points have been placed on straight line $y = 2 \cdot x - 8$. The segment of the line at $4 \leq x \leq 6$ has been considered. Curvature of interpolation line is demonstrated in Figure 14.

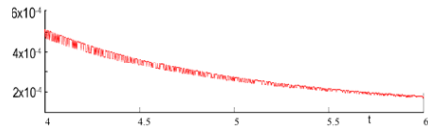


Figure 14: Curvature of interpolation straight line ($\kappa = 0.01$, $t = \varphi$).

Curvature of interpolation line with the same data points and parameter values using Cartesian coordinate system are presented in Figure 15. Curvature comparison is indicative for polar coordinates pattern usefulness.

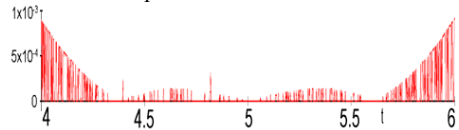


Figure 15: Curvature of interpolation straight line ($\kappa = 0.01$, $t = x$).

Helical line interpolation has been considered too. Its result is indicative for cylindrical coordinate pattern usefulness too. Unexpected results have been obtained at the end points. Extrapolated continuation of curve is close to the test curve.

5. CONCLUSIONS

Continuous interpolation has a geometric smoothness advantage. The interpolation pattern is defined by the initial data and the problem deciding after interpolation.

The explicit and parametric forms of the interpolation function give equivalent results at the uniform or weak non-uniform presented curve. Multiple-valued function demands parametric form of function or its fragmentation. It is possible to use estimate value of arc length as interpolation parameter for non-uniform presented curve. This pattern demands less value of manipulated variable than another pattern.

Parametric presentation is preferable even in 2-D case if data set is represented in non-uniform form.

As expected, qualitative criteria are deteriorated on border regions especially using natural parameterization. Many ways of improvement in the end point accuracy are known.

Requisite smoothness of curve and its characteristics are taken by using correct pattern and manipulated variable.

Polar and cylindrical coordinate systems are both suitable to Cartesian one for interpolation problems.

It is preferable to use the same coordinate system as data points.

REFERENCES

- [1] B. C. de Boor, *A Practical Guide to Splines*, New York, Springer-Verlag, 1978.
- [2] І.Ю. Грищенко. До визначення оптимальної інтерполяції точкового ряду графіками гіперболічного секанса, *Геометричне та комп'ютерне моделювання*, Вип. 28, стор. 62–66. ХДУХТ, Харків, 2011.
- [3] С.Ф. Пилипака, І.Ю. Хименко. Дослідження інтерполяційної функції на основі суми графіків гіперболічного секанса, *Праці Таврійської державної агротехнічної академії*, Вип. 4, Т. 3, стор. 12–15. ТДАТА, Мелітополь, 2003.
- [4] Ю.В. Сидоренко, А.В. Сацкова. Комп'ютерна реалізація різних способів параметризації інтерполяційної функції Гауса, *Прикладна геометрія та інженерна графіка*, Вип. 72, стор. 174–178. КНУБА, Київ, 2003.
- [5] Taras. Data interpolating surfaces using the sum of bell-shaped functions. In *Proceedings of the ICGG (Dresden, Germany, August 4–8)*, 2008.
- [6] І.П. Тарас. Інтерполяція просторових дискретно заданих кривих за допомогою дзвоноподібних функцій. *Міжвузівський збірник за напрямком «Інженерна механіка» (Сучасні проблеми геометричного моделювання, квітень 2008)*, Вип. №22, Ч. 1, стор. 336–340. ЛДТУ, Луцьк, 2008.
- [7] І.П. Тарас, І.В. Павлик. Візуалізація фізичних об'єктів та процесів на основі створення інтерполяційних геометричних моделей. *Розвідка та розробка нафтових і газових родовищ*, 2(35), стор. 83–88. ІФНТУНГ, Івано-Франківськ, 2010.

ABOUT THE AUTHORS

Iryna Taras, Ph.D., is a Associate Professor at the Engineering and Computer Graphics Department at the Ivano-Frankivsk National Technical University of Oil and Gas. Her research interests are Geometrical Approaches in Arts and Design. She can be reached by e-mail: erytar@mail.ru or through postal address: Engineering and Computer Graphics Department, Ivano-Frankivsk National Technical University of Oil and Gas, 15, Karpatska str., Ivano-Frankivsk, 76019, Ukraine.

PATTERNS ON TRIPLY PERIODIC POLYHEDRA

Douglas DUNHAM

University of Minnesota Duluth, USA

ABSTRACT:

This paper discusses patterns on *triply periodic polyhedra*, infinite polyhedra that repeat in three independent directions in Euclidean 3-space. We further require that all the vertices be congruent by a symmetry of the polyhedron, i.e. that they be uniform, and also that each of the faces is a single regular polygon. We believe that we are the first to apply patterns to such polyhedra. The patterns we use are inspired by the Dutch artist M.C. Escher. The patterns are preserved, up to color symmetry, by the symmetries of the polyhedra.

Keywords: periodic polyhedra, Escher designs, hyperbolic geometry

1. INTRODUCTION

A number of artists, including M.C. Escher, have placed patterns on convex polyhedra. In contrast, we show Escher-inspired patterns on *triply periodic polyhedra*, infinite polyhedra that repeat in three independent directions in Euclidean 3-space. Doris Schattschneider's book *M.C. Escher: Visions of Symmetry* is a good reference for Escher's works [3]. We require that those polyhedra (1) be composed of copies of a regular polygon, and (2) that they be *uniform*: all vertices are congruent by a symmetry of the polyhedron. Figure 1 shows a pattern of angular fish on the simplest triply periodic polyhedron, which is based on the cubic lattice.

The triply periodic polyhedra we discuss are often called hyperbolic since the sum of the angles of the polygons at each vertex is greater than 360 degrees (if the angle sum were 360 degrees, the "polyhedron" would be flat; and if it were less than 360 degrees, the polyhedron would be finite).

In 1926 H.S.M. Coxeter and John Petrie discovered a regular class of triply periodic polyhedra which they called *infinite skew polyhedra* and have symmetry groups that are flag-transitive, and thus are natural analogs of the Platonic Solids. They designated those polyhedra



Figure 1: A pattern of angular fish on the $\{4, 6 | 4\}$ polyhedron.

by the extended Schläfli symbol $\{p, q | n\}$, indicating that there are q p -gons around each vertex and n -gonal holes. Coxeter and Petrie showed there are three possibilities: $\{4, 6 | 4\}$ (shown in Figure 1), $\{6, 4 | 4\}$, and $\{6, 6 | 3\}$ [1].

In the following section we discuss the relationship between triply periodic polyhedra and regular hyperbolic tessellations. Then we show patterns on each of the infinite skew polyhedra. Next we show a patterns on less regular polyhe-

dra. Finally we draw conclusions and indicate directions of future work.

2. PERIODIC POLYHEDRA AND REGULAR TESSELLATIONS

We use the Schläfli symbol $\{p, q\}$ to denote the regular tessellation formed by regular p -sided polygons or p -gons with q of them meeting at each vertex. If $(p-2)(q-2) > 4$, $\{p, q\}$ is a tessellation of the hyperbolic plane (otherwise it is Euclidean or spherical). Figure 2 shows the tessellation $\{4, 6\}$ in the Poincaré disk model of hyperbolic geometry. Escher based all four of

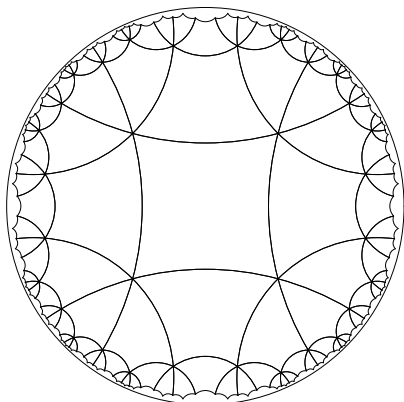


Figure 2: The regular tessellation $\{4, 6\}$.

his hyperbolic “Circle Limit” patterns, and many of his Euclidean patterns on regular tessellations. Figure 3 shows that tessellation superimposed on an Escher-like pattern of angular fish similar to those of Figure 1.

As mentioned, each of the triply periodic polyhedra we consider have a p -gon for each of its faces, with q p -gons around each vertex (since the polyhedron is uniform). Thus we can also use the Schläfli symbol $\{p, q\}$ to refer to such polyhedra, however, unlike regular tessellations, different polyhedra can have the same $\{p, q\}$, as we will see below. We have already introduced the extended Schläfli symbol $\{p, q | n\}$ used by Coxeter and Petrie to (uniquely) specify

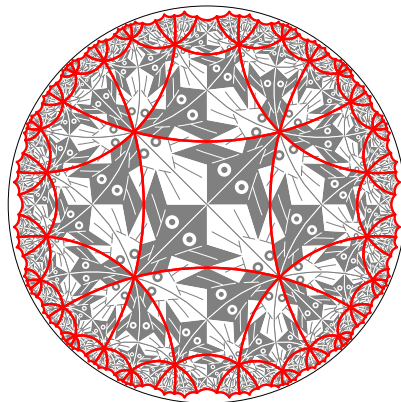


Figure 3: The regular tessellation $\{4, 6\}$ superimposed on a pattern of angular fish.

their more regular infinite skew polyhedra.

For the three infinite skew polyhedra and the other triply periodic polyhedra discussed below, there is an intermediate “connecting surface” between the polyhedron $\{p, q\}$ (or $\{p, q | n\}$) and the corresponding regular tessellations $\{p, q\}$. First, these periodic polyhedra are approximations to triply periodic minimal surfaces (TPMS). Alan Schoen has done extensive investigations into TPMS [2]. Figure 4 shows the Schwarz P-surface, which has the same topology as the polyhedron in Figure 1. In fact the embedded Euclidean lines of the P-surface are just the backbone lines of the fish in Figure 1.

Second, each smooth surface has a *universal covering surface*: a simply connected surface with a covering map onto the original surface, which is a sphere, the Euclidean plane, or the hyperbolic plane. Since each TPMS has negative curvature (except for possible isolated points), its universal covering surface does too, and thus has the same large-scale geometry as the hyperbolic plane. In the same vein, we might call a hyperbolic pattern based on the tessellation $\{p, q\}$ the “universal covering pattern” for the related pattern on the polyhedron $\{p, q\}$. Thus Figure 5 shows the universal covering pattern for the

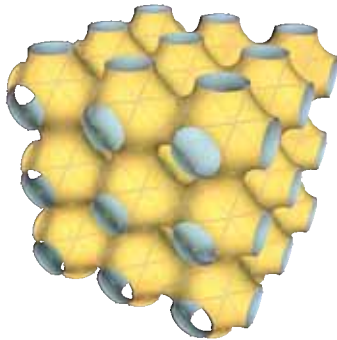


Figure 4: The Schwarz P-surface, showing embedded lines.

patterned polyhedron of Figure 1. In the next section we discuss fish patterns on infinite skew polyhedra.

3. ANGULAR FISH ON INFINITE SKEW POLYHEDRA

In Figure 1 we have shown a pattern of angular fish on the infinite skew polyhedron $\{4, 6 | 4\}$. In this section we also show patterns of angular fish on the other two infinite skew polyhedra: $\{6, 4 | 4\}$ and $\{6, 6 | 3\}$.

The $\{6, 4 | 4\}$ polyhedron is the dual of the $\{4, 6 | 4\}$ polyhedron in which each vertex is replaced by a hexagon. The $\{4, 6 | 4\}$ polyhedron is based on the bi-truncated, cubic, space-filling tessellation by truncated octahedra. The $\{6, 4 | 4\}$ polyhedron divides space into two sets of truncated octahedra which are connected by their square faces. Figure 6 shows the polyhedron decorated with angular fish. Figure 7 shows a top view of the backbone lines of the fish on the $\{6, 4 | 4\}$ polyhedron, which are the same lines as the backbone lines of the fish in Figure 1 — not surprising since the polyhedra are duals. Thus the TPMS corresponding to the $\{6, 4 | 4\}$ polyhedron is the same as for the $\{4, 6 | 4\}$ poly-

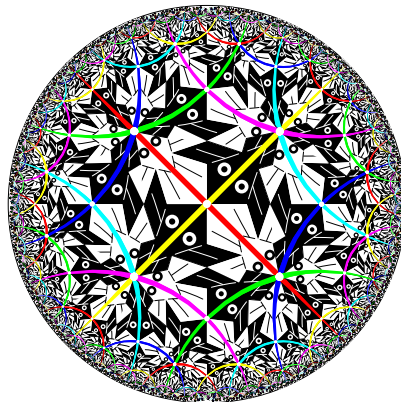


Figure 5: The “universal covering pattern” for Figure 1.



Figure 6: The $\{6, 4 | 4\}$ polyhedron decorated with angular fish.

hedron: the Schwarz P-surface. Figure 8 shows



Figure 7: A top view of the $\{6,4 | 4\}$ polyhedron.

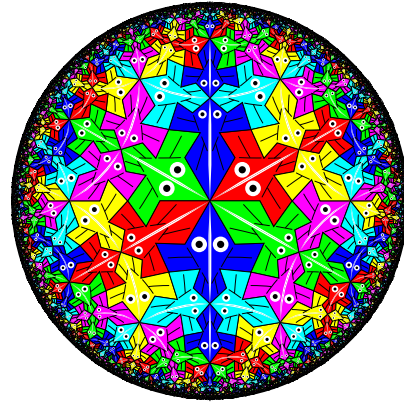


Figure 8: The “universal covering pattern” for Figures 6 and 7.

the universal covering pattern for the pattern shown in Figures 6 and 7.

The self-dual $\{6,6 | 3\}$ polyhedron is more difficult to understand than the other two infinite skew polyhedra. It is formed from truncated tetrahedra with their triangular faces removed. Such “missing” triangular faces from four truncated tetrahedra are then placed in a tetrahedral arrangement (around a small invisible tetrahedron). Figure 9 shows a side view of a $\{6,6 | 3\}$ polyhedron decorated with angular fish. Figure 10 shows a top view looking down at one of the vertices (where six hexagons meet). The corresponding TPMS is Schwarz’s D-surface, where D stands for Diamond [2]. The Schwarz D-surface divides space into two congruent parts, each with the shape of a thickened diamond lattice. It also has embedded Euclidean lines, which correspond to the backbone lines of the fish. Figure 11 shows the corresponding universal covering pattern based on the $\{6,6\}$ tessellation.

In the next section we discuss patterns on $\{3,8\}$ polyhedra.



Figure 9: The $\{6,6 | 3\}$ polyhedron decorated with angular fish.



Figure 10: A top view of a pattern of fish shown in Figure 9.

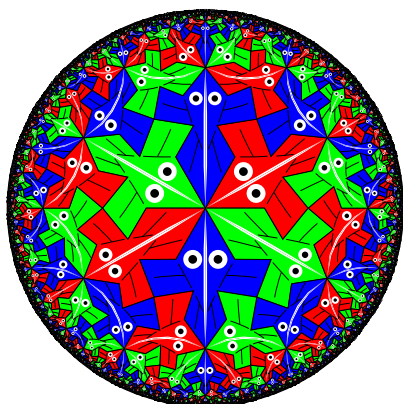


Figure 11: The universal covering pattern for Figures 9 and 10.

4. PATTERNS ON $\{3, 8\}$ POLYHEDRA

In this section we start by examining butterfly patterns on two different $\{3, 8\}$ Polyhedra, so we can see visually that the Schläfli symbol is not enough to specify a triply periodic minimal surface. Then we end by considering a fish pattern on one of those polyhedra.

The butterfly pattern was inspired by Escher's Regular Division Drawing 70, which is shown in Figure 12. Figure 13 shows the correspond-



Figure 12: Escher's Regular Division Drawing 70.

ing hyperbolic universal covering pattern for the patterns on the first two polyhedra that we will discuss.

The first $\{3, 8\}$ polyhedron has the shape of a diamond lattice, like the Schwarz D-surface, and thus the intermediate TPMS is that surface. This polyhedron is made up of parts of regular octahedra of two types: "hub" octahedra and "strut" octahedra. Each hub octahedron has four strut octahedra placed on alternate faces of that hub, so four hub triangles are covered by struts and four are exposed. Each strut connects two hubs to opposite faces of the strut, which are covered by the hubs, leaving six exposed triangle faces. Thus eight equilateral triangles meet at each vertex, and this is a $\{3, 8\}$ polyhedron. However it

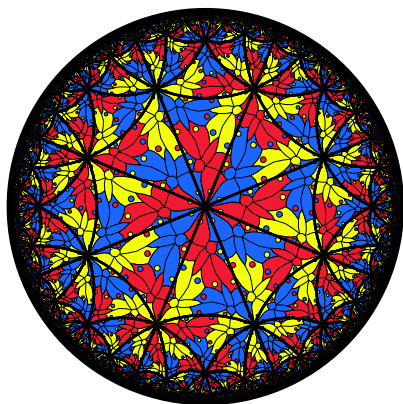


Figure 13: A butterfly pattern based on the $\{3,8\}$ tessellation.

is not as regular as the infinite skew polyhedra since there is no symmetry of the polyhedron that maps a hub triangle face to a strut triangle face (and vice versa). Figure 14 shows this polyhedron covered with butterflies.



Figure 14: Butterflies on a $\{3,8\}$ diamond lattice polyhedron.

The second $\{3,8\}$ polyhedron is composed of skew cubes with alternating chiralities centered at cubic lattice points and connected by their (missing) square faces. Thus the interme-

diate TPMS is the Schwarz P-surface. Figure 15 shows that polyhedron with a butterfly pattern.



Figure 15: Butterflies on a $\{3,8\}$ polyhedron made of snub cubes.

Finally we show a pattern of fish on the “diamond lattice” $\{3,8\}$ polyhedron that was inspired by Escher’s hyperbolic pattern *Circle Limit III*. Figure 16 shows one view of that polyhedron. Figure 17 shows Escher’s *Circle Limit III* pattern with the $\{3,8\}$ tessellation superimposed. Figure 18 is a view of the polyhedron down a 3-fold rotation axis.

5. CONCLUSIONS

We have discussed some of the theory of triply periodic polyhedra. In particular we have concentrated on triply periodic polyhedra that are uniform (all vertices are congruent) and composed of copies of a regular p -sided polygon. We have also shown how those polyhedra can be decorated with patterns that are related to repeating patterns of the hyperbolic plane. In particular, we have shown angular fish patterns on each of the three most regular triply periodic polyhedra, the infinite skew polyhedra of Coxeter and Petrie. We have also shown an Escher-inspired butterfly pattern on two different $\{3,8\}$ polyhedra. Finally we have shown a fish pattern on a $\{3,8\}$ polyhedron that was inspired by Escher’s hyperbolic print *Circle Limit III*.



Figure 16: A fish pattern on the diamond lattice $\{3,8\}$ polyhedron.



Figure 18: The polyhedron showing a 3-fold rotation axis.

In the future we hope to investigate other less regular triply periodic polyhedra. We would also like to place other Escher-inspired patterns on triply periodic polyhedra, either regular or less regular.

ACKNOWLEDGMENTS

I would like to thank the many students who have worked over the years on the software used to create the patterns shown in this paper.

REFERENCES

- [1] H. S. M. Coxeter. Regular Skew Polyhedra in Three and Four Dimensions. *Proc. London Math. Soc.*, (2): 33–62, 43 1937.
- [2] A. Schoen. Triply Periodic Minimal Surfaces (TPMS). <http://schoengeometry.com/e-tpms.html> (accessed 16 June, 2014).
- [3] D. Schattschneider. *M.C. Escher: Visions of Symmetry*. Harry N. Abrams, New York, NY, 2004.

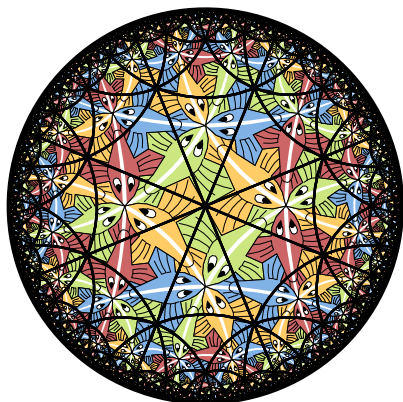


Figure 17: Escher's *Circle Limit III* with $\{3,8\}$ superimposed.

ABOUT THE AUTHOR

1. Douglas Dunham is a Professor in the Department of Computer Science at the University of Minnesota Duluth. His research interests include repeating patterns, hyperbolic geometry, and the art of M.C. Escher. He can be reached at his email address: ddunham@d.umn.edu or via his postal address: Department of Computer Science, University of Minnesota, 1114 Kirby Drive, Duluth, Minnesota, 55812-3036, USA.

PERSPECTIVAL GEOMETRIES IN THE SACRED MOUNT OF OSSUCCIO

Giuseppe AMORUSO¹ Alberto SDEGNO² and Silvia MASSERANO²

¹Politecnico di Milano, Italia ²Università degli Studi di Trieste, Italia

ABSTRACT: The holy path up to the Sanctuary of the Madonna del Soccorso di Ossuccio (Como) is one of the Sacred Mountains of Piedmont and Lombardy, part of the UNESCO site of the nine Sacred Mountains of Northern Italy: a sequence of small religious buildings that host the representations of the Mysteries of the Rosary, in a hybrid of sculpture and painting, an illusory set which melts with architecture's geometry. The morphological configuration and the figurative contents of each chapel require a specific analysis addressing the four thematic layers of geometry, perspective, painted architecture and landscape. The buildings stress the relationship with the landscape in an original interpretation of perspective-based composition and painted decoration (the baroque *quadratura*), which conceals the shapes of the chapels' interior. The aim of this study, part of a national search started in 2013 about illusory perspective in architecture, identifies the workflow to describe the perspectival geometries of the chapel entitled *The Jesus' dispute with the doctors in the Temple*, an application of painted perspective to the octagonal shape of the building to improve the perception of the space.

Keywords: Geometry, perspective, *quadratura*, drawing, 3D modeling.

1. THE SACRED GEOMETRY

The study presented here is part of a research about the architectural perspective and it is investigating the geometry of a series of chapels located on the Sacro Monte di Ossuccio in the Italian province of Como.

Their particular morphological configuration and their figurative contents require a careful analysis to reconstruct the three levels at the base of the research, that are architecture, painting and sculpture. Within each chapel, in fact, you can find wall paintings and statues in real scale that evoke the mysteries of the Rosary. The aim of this study was to identify the operating stratagems to capture and describe those works which, for their specificity, are part of a national search started in 2013 about the building of illusory perspectives and the architectural space.

In fact, the 14 chapels of the Sacro Monte in Ossuccio are each a small and popular *Gesamtkunstwerk*, that is a total work of art in which architecture, painting and sculpture fully

participate in the unfolding of the description. Built during the Seventeenth Century, the holy path to the Sanctuary of the Madonna del Soccorso di Ossuccio (Como) is one of the Sacred Mountains of Piedmont and Lombardy, since 2003 part of the UNESCO site of the nine Sacred Mountains of Northern Italy [1].

They are designed as a sequence of small religious buildings, with a central plan – circular, square or octagonal –, that host the representations of the Mysteries of the Rosary, in a hybrid of sculpture and painting, an illusory set which melts with architecture's geometry. The morphological configuration and the figurative contents of each chapel require a specific analysis addressing four thematic layers: geometry, perspective, painted architecture and landscape. Buildings stress the relationship with the landscape in an original interpretation of perspective-based composition and painted decoration (the baroque *quadratura*), which conceals the shapes of the chapels' interior.

Chapels are quite similar in size and architectural morphology, and also contain analo-

gous figurative equipment although everyone evokes a different mystery having a volume that does not exceeds 1000 cubic meters. The chapel described in this paper is the one entitled *The Jesus' dispute with the doctors in the Temple* [2–3], the fifth of the series: the space is based on the octagonal geometry and the entrance is featured by a *pronaos*; the distance between major sides of the polygon is about 9 m and height is about 10 m. The building was erected by Giovanni Somalvico in the period 1683 – 1688 and it has been replicated several times, even if with variations, in other chapels located along the same votive path.

The aim of the graphical investigations was to identify the workflow to describe the perspectival geometries of the chapel, a specific case based on the application of painted and illusory perspective to the geometric shape of buildings to improve the perception of the space. The geometric documentation considered the formal qualities of the chapel according to its figurative design. The chapel is composed of some quadrature painted on planar walls and also a perspective painted on the vault that accentuates the spatiality illusion of its decorative elements.



Figure 1: View of the interior space of the Chapel. In the middle the sculpture of Christ.

The painted ceiling in the center of the vault shows the columns in perspective view and it is explicitly marking the vanishing point according to the observer who turns his gaze visiting the chapel from the entrance door. (Fig. 1)

These small architectures are settled in a

very natural context and the particular location influences the survey process because of several obstacles and trees. The evolution of survey tools and techniques makes available various and not homogeneous technical equipments: from more precise hardware – such as 3D scanners or cameras with different qualities and accuracies – to photo modeling software – that allows the semi-automatic treatment of three-dimensional data and raster images. Although developed for a generic operability – that should satisfy most of the demand coming from the survey market – it is still possible to identify standards, limitations, where the strict application of a procedure could require further intervention methods that have not yet been studied, to be included within the scope of an experimental research in the field of representation.

With regard to the importance of the interior space, due to the reduced size of the chapel and to the density of decoration, you have to deal with blind spots that hinder the full documentation of the details that are not easy to access. Moreover, the presence of groups of sculptures, which describe the specific theme of the chapel, would require a multiplicity of scans in order to obtain a high quality of both architectural elements and statuary, suggesting the use of an optical triangulation scanner for the acquisition range (between 50 and 200 cm).

Further limitations are related to the exterior where the architecture is combined with the environmental position: in fact they are located on the edge of the sacred paths and almost always in the middle of dense vegetation that limits the operation of the 3D scanner and reduces the possibility to document the exterior walls in their formal integrity. Further limitation regarding the relief of the external walls, derives from the fact that the architectures stay along a path clearly delineated and that therefore the technical instruments must necessarily rest on the walkway, which preclude the visibility of some parts of the building (poor accessibility to the backs but often also to the side,

due to the perspective path of the perceptive system, which binds the visual to frontal views). We must also point out, however, that, for what concerns the aim of the research currently in progress, is not particularly significant the metric data of the survey, and in particular the one relating to the external volume, for what is sufficient a typological analysis. We therefore focused on the acquisition of point and photographic data of the interior space, with the aim of identifying the operative stratagems for analysis, documentation and communication of architecture, paintings and sculptures [4-5-6].

2. THE GRAPHIC DOCUMENTATION OF THE ILLUSORY ARCHITECTURE

The 3D-modeling methodologies are focused on the acquisition of three-dimensional geometry, with less accuracy than the one that is potentially achieved for small objects and close-range scanners or with photogrammetry tools and calibrated cameras, but equally effective for the purposes of the typological and specific interpretation of the chapels.

Tools and techniques assessment considered a series of factors according to the quality of data and the degree of detail starting from the final 3D model. Briefly these criteria were: the site features, designed as a theatre stage with figures (statues) and painted surfaces as scenes; the maximum overall dimensions and other features related to the different typologies of architectural surfaces and decorative elements; the minimum detail and the artistic technique; type and condition of materials, (considering surfaces and materials as Lambertian); sensors features, photographic camera and laser scanner operating conditions and 3D modeling software; site conditions and accessibility; the overall aims of the survey and following steps.

For what concerns the instruments of acquisition it was deliberately used a set of mid-range products, with good performance, excellent flexibility and versatility during the work on site and especially with a low cost profile, considering the best practice framework

and further applications for cultural heritage and digital archives of architecture. According to this statement, the choice was a Faro Focus 3D laser scanner with a digital camera Nikon Coolpix P7000 to increase the quality of photographic textures during post-production and orthophoto mapping onto the three-dimensional geometric model.

The Faro Focus 3D sensor belongs to the category of the optical-mechanical devices that measure the spatial position of points belonging to surfaces with particular physical characteristics; this tool belongs to the category of phase measurement active sensors and measures the phase shift between the emitted signal and its return phase through the emission of optical radiation with a wavelength corresponding to the infrared $\lambda=0.78$ mm. The surface material found in the chapels – groups of terracotta statues and frescoes and painted walls currently in good condition and free from chalking, in the dry state – can be considered as surfaces with a Lambertian behavior. We can estimate for it, in theory, a value of high level reflectivity (ρ), equal to 94% , where the above parameter is valid for a laser with a wavelength of ~ 0.9 microns. According to these operating situations and depending on the shooting distance, accuracies of the order of a few millimeters can be achieved. The main features of Faro sensors are the high-speed shooting and the measurement rate up to 976,000 points/sec; the distance accuracy up to ± 2 mm, with a range from 0.6 m to 120 m and the small and compact size (the smallest 3D laser scanner ever built and weighing only 5 kg/11lb).

The Chapel V, whom interiors are rich of formal, architectural and decorative issues, was surveyed with a set of 4 recording stations settled inside, 2 further stations at the entrance and 3 outside; the external surfaces and the public path, as we mentioned above, were considered less important for the research topics. It was also planned to set the instrument in a range of about one meter from the main sculptural groups reducing the number of the scans and

achieving a most efficient method and clouds density for 3D measurement of geometry (Fig. 2). The 3D documentation also considered the formal qualities of the chapel according to their figurative design: the Chapel V is composed of some quadrature painted on planar walls and there is also on the vault a central perspective that accentuates the spatiality illusion of its decorative elements.

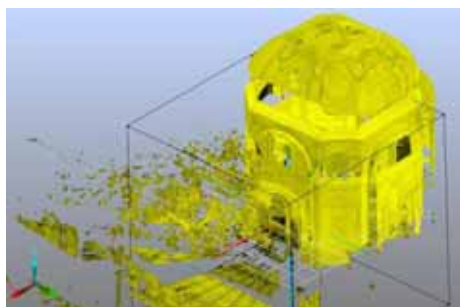


Figure 2: The points' cloud of the chapel.

The frescoes in the Chapel V, however, do not show at first sight inconsistencies, except for the lack of a capital and of the color layer of a pilaster painted on the left side entrance designed as a *serliana* door (Fig. 3), decorative elements that are clearly visible in the opposite entrance door on the right side (Fig. 4).



Figure 3: The painted wall with the Serliana on the left side of the chapel.

We also underline the correct representation of the ceiling in the painting in the center of the vault, which shows the columns in perspective

view (Fig. 5) explicitly marking the vanishing point according to the observer who turns his gaze visiting the chapel from the entrance door.



Figure 4: The painted wall with the Serliana on the right side of the chapel.

During the following step, cloud registration and 3D modelling it will be possible to verify accurately consistencies and inconsistencies of paintings and their subjects.



Figure 5: The columns in perspective painted on the vault.

With regard to the groups of sculptures—modeled in clay and painted with gouache by Agostino Silva—the chapel presents the figure of the young Jesus, placed in the center standing with the halo, which is talking with the doctors in the Temple of Jerusalem. Twenty human figures, in this case, are placed in the room, from which you can identify the Virgin Mary and the Apostle Peter, as well as a book and a dog sitting, both of them in the middle of the room. In the background, behind the sculptures, there are other subjects painted on the walls, which appear between the painted Ionic

columns. Much richer than other chapels, referring to the iconographic aspect, the survey process required greater care in positioning the scanner.

After completing the scanning phase, the point clouds (range map) were registered through a process of alignment, using the Faro Scene modeling software. In this first phase was essential the process of noise reduction and data refinement, which allowed us to eliminate the instrumental noise that could create false surface definition not belonging to the scanned scene. The data correction is more important if using acquisition tools that have a level of accuracy very high and therefore generate point clouds having more geometric inconsistencies.

After registering clouds and making alignment, the model was visualized both in a subjective way, with perspective views of interior spaces to simulate the behavior of a panoramic projection, and in objective way, with views from outside of the buildings depicted in isometric projection.

This first survey framework has to be classified as experimental testing for the development of an operating procedure; further developments will be the definition of a process of integrated survey, employing techniques of digital photogrammetry and verifying the degree of integration of image-based 3D modeling, considering also the high-resolution photography.

3. THE PERSPECTIVE ANALYSIS

The investigation started from the analysis of the zenithal perspective. The convergence of the perpendicular lines to the projection plane (identified in the edges of the pillars supporting the three arches) has determined the position of the principal point - and the relative point on the floor - on the octagonal frame of the vault. The location of that vanishing point should have ensured an easy control of the perspective heights also in the realization of the work. Then, we considered the painted perspective on the left side of the wall, from the entrance of the

chapel. In this case the projecting lines, similar to straight lines that are orthogonal to the plane were identified as directions of the moldings of the architrave below the arch of the serliana. The convergence of these perspective lines belongs to the edge of the white pilaster and is located at the height of 2.47 meters from the floor. We found that the position of this principal point is perfectly aligned with the main station located on the painted perspective on the vault. Although this vanishing point confirms the highest number of references it does not meet the perspective orientation of the plane of the arch that put the horizon in a lower position (Fig. 6).

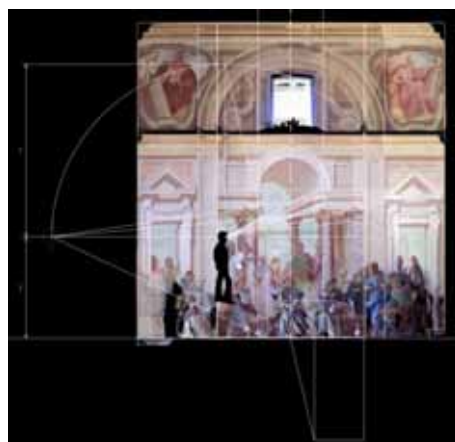


Figure 6: The graphic analysis and the perspective restitution of the Serliana.

It should be noted that the architectures represented in the background are related to another perspective system.

To determine the distance we considered the main geometric construction of an element belonging to the architrave moldings. Accepting the hypothesis that the plan alignment angles drawn by the moldings follows a direction in-

clined of 45 degrees to the plane of the facade, we identified, on the horizon traced in correspondence of the main point, the distance points and we quantified the main distance with the measure of 4.20 meters.

The same construction has been carried out for the wall at the right of the chapel. Also in this perspective, the horizon is at 2.47 meters from the floor, while the main distance has a dimension of 4.23 meters.

As considered in the previous perspective system, the main point does not confirm the total confluence of projecting lines: from this kind of system the convergences of the plane of the arch and the various depths relative to the architectures in the background are in fact excluded.

The geometrical analysis of the main perspective, realized in coherence with the previous hypothesis, confirmed for the main point of this apparatus a central position situated at a height of 1.62 meters from the floor of the choir. The measure is compatible with the eye of an observer, having an height of about 1.75 meters. As previously noted, not every projecting lines converge at the same point: in this case some moldings painted on the right remain excluded. Probably it is an adjustment occurred in the course of work, due to the asymmetric dimension of the projection of the right frame, significantly greater than the one represented on the left. The main distance was instead quantified in 5.31 meters.

Contextualizing in plan these results we specified the best point for the perspective system. In this context, the observation point is almost physically occupied by the sculpture that depicts the Ghignarello – a particular subject who represents a real man visiting the place during the realization of the fresco –, which seems to indicate in which direction you have to turn your eyes to grasp the main scene. Its

finger is crossed by the optical axis identified by the location determined.

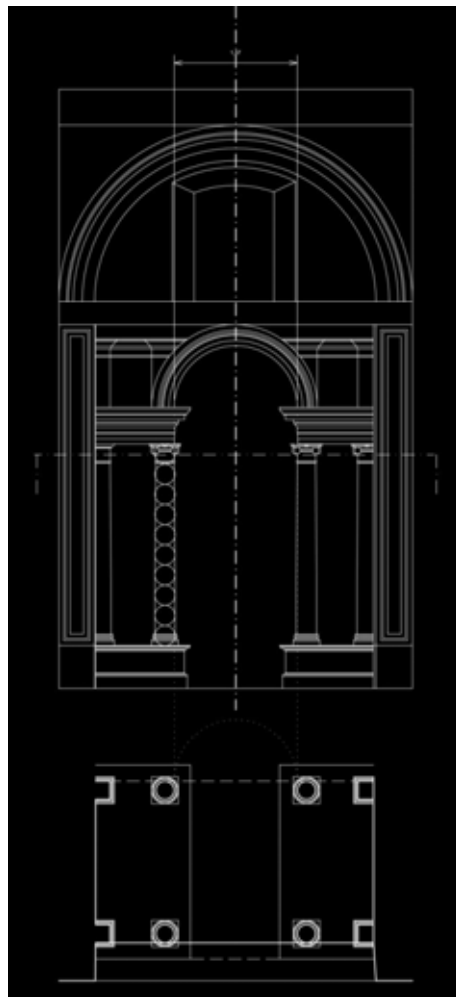


Figure 7: The painted wall with the Serliana door on the left side of the chapel.

This figure, realized as the others in terracotta, is clearly visible from outside, because it does not obstruct the view as is sitting on the

step. Although the whole perspective system is constructed in function of a spectator, the chapel has in façade a triptych of openings to be observed of the interior space by the pilgrims. If you put an observer outside the chapel and near of the center door, you can notice that the railing above the height shows some elements which are enriched by geometric designs, which limit the internal view to direct it in the desired direction. Then, from outside the main perspective is perceived without distortion as documented by the amplitude of the optical cone drawn in plan considering the physical limits of the central door.

We can reflect, however, on the reasons of two different horizons, the first associated to the perspective in front of the entrance, the second one associated to the others, in an higher position in reference to an human eye. Probably the choice is due to the presence of human-sized sculptures, that are arranged in front of the walls, hiding a considerable portion of the painted surface, and forcing the painter to put his characters to a higher level in order to make them visible, but without representing them from the bottom to above, risking to make them not recognizable.

Starting from these considerations we drew the elevation and the plan of the Serliana in order to understand the proportions of it inside the space of the chapel. (Fig. 7).

4. CONCLUSIONS

The paper will show the research developments and the advantages in using high-end 3D digital models compared to photographic methods as the visualization of the fruition paths and the visual cones that were designed to let pilgrims to have the whole fruition of the chapel: a spatial orientation first and then the symbolic immersion in the small universe of the painted space. Besides the reconstruction allows to include the traditional perspective analysis of the quadratura giving the possibility to address the scientific and artistic manufacturing process of

these illusory spaces and to simulate the user approach as a dynamic discover fruition of the illusory architecture. The geometrical analysis of the perspective allowed us to understand better the space of the chapel, and the contribution to the story of the quadratura by the artists involved in this art work. The next phase will be the digital reconstruction of the 3D model related to the perspective representation, and the simulation inside the virtual scene.

ACKNOWLEDGMENTS

This study is part of the Prin 2010-11 Italian National Research program titled *Architectural Perspectives: digital preservation, content access and analytics* (General Investigator: Riccardo Migliari) and funded by the MIUR, University Ministry. The material presented here was developed within the Research Unit of the Politecnico di Milano coordinated by Michela Rossi. Giuseppe Amoruso, Vita Maria Firenze, Piero Lusuardi operated the field survey and the first data analysis; Carlo Alberto Bozzi registered data and edited the orthophoto extraction, Alberto Sdegno and Silvia Masserano provided the perspective analysis and graphic restitution.

REFERENCES

- [1] F. Fontana and Sorrenti P. Sacri monti: note architettonico-urbanistiche. In *Atti del Convegno Internazionale sui Sacri Monti (Varallo Sesia, 1980)*, Tipografia Editrice Zanfa, Varallo, page 210, 1980.
- [2] SIRBeC Lombardia Region Beni Culturali, form ARL n. CO250-00013. (<http://www.lombardiabeniculturali.it/architettura/schede/CO250-00013/>, last access June 2014).
- [3] P. Gatta Papavassiliou. Il Sacro Monte di Ossuccio. Guida alle Cappelle, G. Mondadori, Gorle (Bergamo), page 188, 1996.
- [4] G. Amoruso and L. Galloni. Interactive Environments in Cultural Heritage. Interiors and construction of San Francesco church in Bologna. In *Proceedings of the Interna-*

tional Congress "Domes in the World". Nardini Editore, Firenze, 2012.

- [5] A. Sdegno. RQS - Reverse Quadratura for Surveying. Applying Ancient Methods to Digital Techniques. In *Proceedings of the XXV eCAADe Conference (Frankfurt am Main, Germany, September 26-29)*, pages 615-620. Kanne Graphischer Betreib GmbH, Ginsheim-Gustavsburg, 2007.
- [6] A. Sdegno. Didactic Experiences of Digital Representation of Architecture from 3D Survey to 3D Simulation. In *Proceedings of 2010 ENHSA-EAAE Conference (Istanbul, Turkey, June 17-19)*, pages 371-380. Yildiz Technical University, Istanbul, 2010.

spective analysis and in the applications of Descriptive Geometry in Architecture.

ABOUT THE AUTHORS

1. Giuseppe Amoroso is assistant professor at Politecnico di Milano School of Design and Doctor of Philosophy in Drawing and Measured Drawing (2001, Polytechnic University of Marche). He edited several essays on architectural measured drawing, historic centers documentation, computer graphics and representation techniques. He is Chair of Intbau Italia, the Italian chapter of the International Network for Traditional Building Architecture & Urbanism of London.

2. Alberto Sdegno is assistant professor at Department of Engineering and Architecture at the University of Trieste, and PhD in Survey and Representation in Architecture and Environment at the University of Naples "Federico II". Its main researches are developed in the field of representation of architecture, digital drawing, advanced survey, new technology of visualization.

3. Silvia Masserano is PhD student at the Department of Engineering and Architecture at the University of Trieste. She taught Descriptive Geometry and Representation of Architecture as contract professor in the same school. Her main researches are developed in the per-

PERSPECTIVE BETWEEN FICTION AND FUNCTION: PATTERN MUTATIONS THROUGH SCIENCE AND ART

Luigi COCCHIARELLA
Politecnico di Milano, Italy

ABSTRACT: This paper aims to focus on the *mutations* occurred in the perspective pattern as signs of changes in the projective knowledge and in the same idea of space. Nowadays a Perspective manual normally starts showing ‘the end of the story’: an adamantine reference system made of points, lines, circles and subtended planes, namely main point and main distance, horizon and ground lines, distance and visual circles (Fig. 1b). As well as a Projective Geometry manual does (Fig. 1c). They both are significantly different from that described at ‘the beginning of the story’ by Leon Battista Alberti in *De Pictura proestantissima et numquam satis laudata*, the first official treatise following the Brunelleschi’s perspective code: a point and a line, namely the main point and a segment on the ground line, with a separate schematic side view of the system, superbly graphically combined by Piero Della Francesca in the book *De Prospectiva Pingendi* (Fig. 1a).

About five centuries elapsed between the first and the last scheme in the Figure 1, and although the three systems work as operational *menus* in three graphic ‘releases’ of the same projective ‘software’, it might be hard to believe it by simply comparing the mentioned diagrams. In fact, the Renaissance method assumes a not-extended Euclidean space, still wanting the ideal elements, where only the centric point is handled; in technical terms, the receding elements are located according to a foreshortened squared ground plane, distorted in advance by using the intersection point appearing in the overlapping side view. The reference system is still affected by the *ballistic principle* generating the perspective method, and referred to a confined space. We will call it *oriented projective box*. On the contrary, the current model lays on the hypothesis of an extended Euclidean space including the ideal elements, whose projections as vanishing elements furnish the key bases for the graphic representation; in technical terms it uses the homology as a key algorithm to relate figurative and metrical information, and all the constructions can be drawn on the picture plane, at the point that the reference system itself is set-up as a *homological system*. The horizon and the ground lines humanize this mathematic scheme, which shows an *oriented projective space*. While the Central Projections set-up is completely independent on any human orientation, showing a pure *extended projective space*.

The abovementioned cases reveal significant differences in terms of geometric approach and graphic procedures: as Robin Evans could say, the earlier is the method of Art, aiming to provide affordable constructions to the pictorial *fiction*, while the others concern the method of Science, focusing on the projective *function*. Therefore, a question about ‘what’ the two systems actually show would emerge: according to the spirit of the ages, starting from showing the *image of the objects*, it seems that perspective has gradually shown the *image of the space itself*.

Keywords: perspective, homology, projective geometry, descriptive geometry, history of geometry.

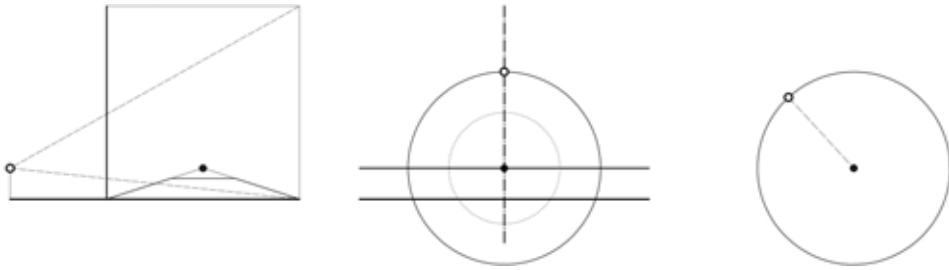


Figure 1: Perspective pattern mutations

a) left: early setup, XV Century (Piero Della Francesca), or the *oriented projective box*, b) midst: present setup, XVII-XIX Century, or the *oriented projective space*, c) right: central projection setup, XIX-XX Century, or the *extended projective space*. Drawn by the author.

1. THE DIATRIBE: “FIGURE” AND “MEASURE”

When speaking about *perspective*, people normally get fascinated by the beauty and by the visual realism of the images. Quite differently from when dealing with *axonomic* and *orthogonal projections*, the aesthetic feeling tends to overcome the metric sense in their perception. In fact, over centuries we have been largely used to relating perspective to Art, and the other mentioned representative forms to Science and Technique. In this paper, we would like to remind that this is just a familiar habit. The fact is that since the beginning perspective, from which all the other forms of representation derive, aimed to work on the relationships between the geometry of images and the geometry of space. It grew up relating the Euclidean principles of Optics with the Euclidean spatial knowledge based on the book of the Elements (III Century BC). The intrinsic connection between *figure* and *measure*, that is to say between *graphics* and *geometry*, should clearly arise from the compared schemes proposed by Filippo Camerota (Fig.2)[2]. We know that Filippo Brunelleschi, the author of the first *perspective demonstrations* (ca. 1410), better known as the father of perspective, was also expert in the field of survey. In fact, the principle of *projection/section* he adopted, as described into the Leon Battista Alberti's treatise *De Pictura* (1435), can be used either for taking measures or for taking or generating images.

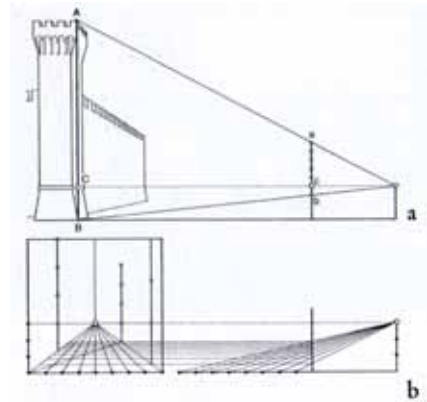


Figure 2: The double-dealing mode of the projection/section process: ‘measuring’ mode (above) and ‘imaging’ mode (below), according to Filippo Camerota [2].

Other two reasoning can be brought forward at this point. The first one concerns the wide use of perspective in the Renaissance painting. Compared to Science, Art appeared more advanced at that time, as painting seemed to be the most effective way to fix/produce ‘verity’ through a human product, even more true thanks to the use of proved geometric processes. As Erwin Panofsky stated [12], perspective allowed an “objectification of the subjectivity”. In other words we have good reasons to think that, as well from a theoretical point of view

Science was still incorporated in Philosophy, as from an experimental point of view Science was still embedded in Art. The second issue concerns the fact that, stressing the image of the real space by sometimes very strong projective distortions, Projective Geometry itself stimulated studies and research about new spaces and geometries. At the point that Arthur Cayley (1821-1895) stated that Projective Geometry itself would be *the* Geometry, being it able to include and to manage the properties of all the other geometric branches as well.

2. EARLY STAGE: BALLISTIC SETUP

As space has three dimensions against the two of the picture plane, since the beginning the drawn setup had to take advantage of at least two preparatory views, including at the same time the location of the viewpoint, of the picture plane and the feature of the real object to represent. Width and height could be directly and easily managed in a front view coinciding with on the iconic plane, while the foreshortenings needed to be found with the help of a side view drawn on a graphic space apart, as Leon Battista Alberti recommended, which did not work in favor of the purity of the graphic constructions, as some measures had to be taken and manually translated on the perspective canvas. This obstacle might have suggested to combine the two views in one, as amazingly represented in the figure XIII by Piero Della Francesca, where the side view is finally rotated on the picture plane (Fig.3). The wedding between Optics and Elements was finally celebrated and every geometrical configuration could be easily shown in perspective. Using a modern language, *metric* and *position* became strictly related. On our opinion, this was the early crying towards the homology.

In *De Prospectiva Pingendi* (c.1475) Piero Della Francesca proposes other two methods for drawing perspective images of any possible shape by using, in different ways, auxiliary orthogonal projections (*ante litteram*), the first one replaying the *ballistic* process carried on in the true space by using a wire, as shown in the famous images of the Albrecht Dürer's treatise

Unterweysung Der Messung (1525). This part includes the most beautiful drawings of the Piero's treatise, but since we decided to follow the homology development, we will not talk about these methods. For the same reason, other practical procedures like those using predefined foreshortened grids on the picture plane, as in the famous Paolo Uccello's (c.1450) *sinopite*, will not be considered as well, in spite of their wide diffusion over the European area.



Figure 3. Piero Della Francesca's figure XIII. Perspective setup combining front view, side view, and perspective image (c.1475) [21].

Then, going back to our story, once defined the perspective setup, Piero shows us a new surprising image in the Figure XVa, where the perspective of a tiled ground is related to its true size representation, supposed to be rotated on the picture plane (Fig.4). In spite of the absence of the viewing distance, this drawing can be considered as the first 'ancestor' of a homological representation. Only deep lines are foreshortened towards the main point, while a diagonal is here presented, that will be used as an auxiliary constructive element in the following examples. Even more surprising is the Figure XXIII, where the diagonal is extended outside the image of the foreshortened ground, to the line of sight. At a first glance it works as the vanishing point of the diagonal, but differently from other authors, Piero does not use it for all the diagonal lines, and most of all he does not use it again in the treatise.

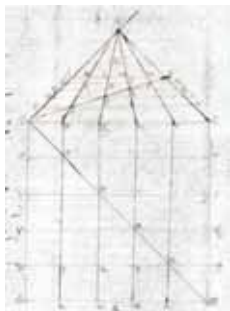


Figure 4: Fig.XVa. Perspective and true shape of a tiled ground, graphically related[21].

This is the reason why some theoreticians, like Giusta Nicco Fasola [6], think that it has been used for the first time as a properly said *distance point*, allowing to directly showing the viewing distance on the perspective drawing, as the distance of that point from the main point. It is clear that this construction has a theoretical value, recommendable but not strictly necessary to realize a convincing perspective, as the viewing distance can be quite freely chosen. Therefore, Piero prefers to discuss many practical examples, referring them to the perspective of a square outline, in other words working into an ideal perspective box.



Figure 5: Fig.XXIII. Perspective of a long corridor subdivided into squared areas by using the construction of the distance point [21].

The double function of the viewing distance, as a way to marking receding points on the canvas, and to finding the perspective meeting point of pencils of diagonal lines, is clear in the Jacopo Barozzi Da Vignola's *seconda regola*.

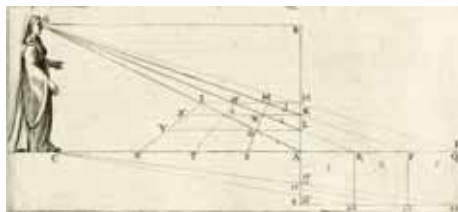


Figure 6: The *seconda regola* by Jacopo Barozzi da Vignola (1583). Double perspective function of the viewing distance [18].

This diagram, coming one century later, can be considered as a further development of the Piero's setup, either because of the insertion of the top view, or due to the use of the distance point as a meeting point of perspective diagonals. For decades, this point has been used as the only auxiliary point in the perspective constructions, not only in relation to horizontal but also to the vertical faces of the objects (Fig.7).

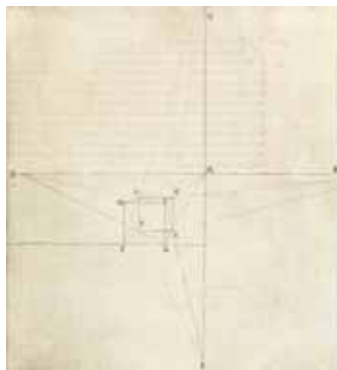


Figure 7: Extension of the notion of distance point to the vertical faces of a cube represented in perspective, from the *Commentarii* of Egnatio Danti to the Vignola's treatise [18].

3. WORK IN PROGRESS: THE NEW VISION

Between XVI and XVIII centuries, when mathematicians became more interested in perspective, the focus gradually changed, moving from the represented object to the representational structures. The Guidubaldo Del Monte's treatise titled *Perspectivae libri sex* (1600) shows important advancements in this field. The demonstrations are clear and very well understandable. Before showing the perspective image of specific objects, a method to represent in perspective any lines and planes however oriented in the space is shown, as clearly anticipated in the drawing printed on the frontispiece of the book (Fig.8). The treatise also includes a chapter devoted to the *scenes*, which means, how to realize three-dimensional perspectives for theatrical representations.

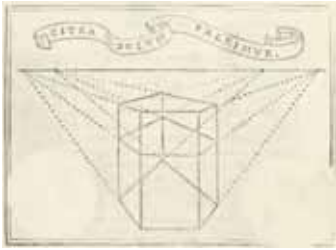


Figure 8: The *Perspectivae Libri Sex* (1600). Frontispiece showing several concurring bundles of horizontal lines [19].

As the properly said notion of vanishing point was still unknown because of the lack of the notion of point at infinity, Guidubaldo highlights the projective relationship between true directions of bundle of lines in the space and meeting points of the corresponding perspective pencils of lines into the image. These meeting points are called *puncta concursus*, that is, concurrent points, what we nowadays call vanishing points. At the end of the sixth book an image is proposed where apart from the vertical lines no other lines are parallel to the picture plane, and the lines belonging to the

sloping ground plane and all the other sloping lines parallel to them, meet at a concurrent point located above the horizon (Fig.9).

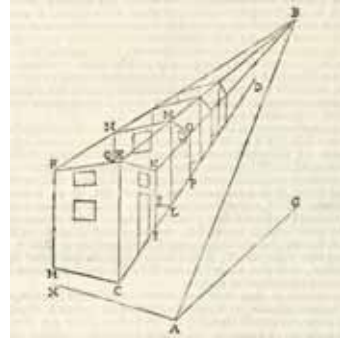


Figure 9: *Perspectivae Libri Sex*. Perspective representation including bearing and sloping bundles of lines [19].

To make the constructions as clear as possible, in the first book not only the graphic procedure on the picture plane, but also *how* these procedures are generated according to the spatial location of the elements, it is carefully shown, what makes this treatise definitely modern.

Anyway, once the connections between space and image have been explained, a graphic method to draw perspectives is afforded. At the early beginning of the second book he proposes a *rabatment* method where either the true figure or the viewing distance, or the sight height and the visual rays are rotated on the picture plane and used for the construction of the perspective images, that of course appear on the same picture plane.

As an example, we have selected the construction XXIII from the fourth volume of the treatise (Fig.10). As we see, the true shape and its perspective appear inverted by the effect of the rotation about the ground line, below which, together with the true figure, in this case a circle, the sight point and the visual rays used to finding the concurrent points are rotated as well. To find the right directions of the visual rays, the center of the circle H and the two lines HI

and HZ need to be rotated, beforehand, above the ground line; for the sake of simplicity, in the figure only the point T, corresponding of H, is drawn above that line.

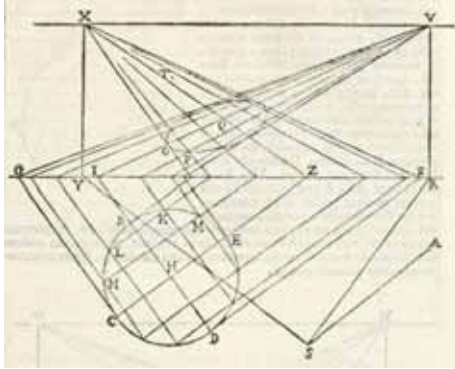


Figure 10: Perspective of a circle, basing on two orthogonal directions, generically oriented in the ground plane [19].

Compared to the *rabatment* proposed by Piero Della Francesca, the metrical information about the sight point are here explicit. However, although totally carried on into the picture plane, the constructions show some discontinuities. Firstly, the abovementioned procedure for the recognition of designated directions, needing a double spatial rotation of a geometrical element. Moreover, the horizon plane is not used to find out but only to show the meeting points of perspective parallel lines: in fact, the visual rays are drawn on the ground plane, starting from the *foot* of the observer (point S) instead of from its *eye* (point A); consequently, the meeting points X and V are not directly defined as points of the horizon plane, but by translating on the horizon line the points Y and R, previously defined on the ground line. But the spatial setup was finally clear, and in the last part of the treatise devoted to the scenography, the new knowledge and constructions are also applied to any possible location of elements in the space, including vertical planes (Fig.11).

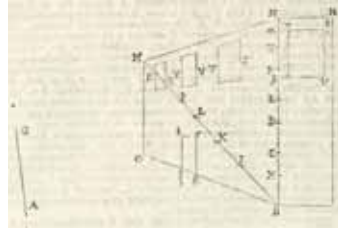


Figure 11: Perspective construction of a vertical plane in a scenic space [19].

Published in 1600, the Guidubaldo's treatise marks the threshold between two centuries and eras. All in all, it represents a great step towards the modern projective approach and, last but not least, towards the discovery of the projective links between true shapes and foreshortened images, that is, the basic ingredients of the homology. In the following decades of XVII century the theory of perspective could also take advantage of the new scientific ideas.

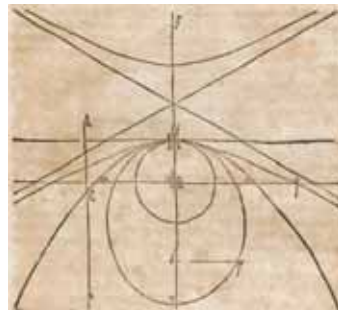


Figure 12: Johannes Kepler, *de coni sectionibus* (1604) [20].

A revolutionary idea was proposed in the book *Ad Vitellionem Paralipomena* by Johannes Kepler (1604) who finally clearly defined the notion of *point at infinity* (Fig.12). He found it in relation to Optics, while discussing about the second focus of the parabola, which is, as we nowadays now, infinitely far. He also stated, for the first time, that parallel lines meet at the point at infinity, so opening the way to the modern idea of vanishing point, intended as a

projective image of a point at infinity. Some years later, Girard Desargues proposed new revolutionary ideas, like the projective analogy between conic and cylindrical projections, and the projective equivalency between straight lines, circle, degenerate conics. From a strictly projective point of view, his theorem on the corresponding triangles, appeared in the *Livrette de perspective adressé aux le théoriciens* (1643, lost), draws the very fundamentals of the homology.

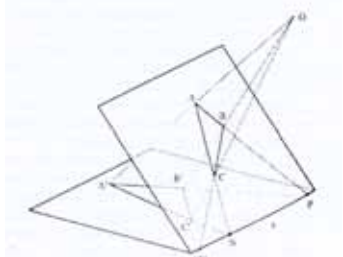


Figure 13: Girard Desargues (1643). Corresponding triangles, spatial version (scheme)[5].

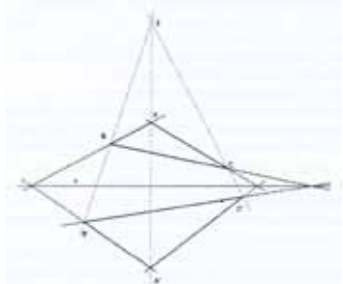


Figure 14: Girard Desargues (1643). Corresponding triangle, plane version (scheme)[5].

The spatial configuration resumed in the figure 13 shows the basic elements of a perspectivity between two planes, namely center, axis, corresponding elements. The plane scheme in the figure 14 shows, at the same time, a perspectivity between two figures belonging to the same plane (or to two overlapping planes), and, if considered as a perspective projection of the configuration in the figure 13, a homology, in spite of the fact that this name was used only

almost two centuries later, by Jean Victor Poncelet. Other important properties of the space and of the conics were discovered and proposed by Desargues by using the graphic methods of perspective, that from that moment on became to acquire an independent theoretical value, so preparing the way to the birth of Projective Geometry. As stated by Jean-Pierre Le Goff and referred by Anna Sgrossso [6], this new use of the projection shows one of the early heuristic applications of perspective to Geometry. In other words, if in the past geometry was used to enhance perspective, from this moment on also perspective could be used to enhance geometry. Anyway, a new theoretical sensitivity was now diffused, and a new approach to perspective can be found that takes into account the new geometrical ideas, not only into the literary masterpieces, but also into some minor work. Among these, we propose in the figure 15, from the treatise *Perspective speculative et pratique* by Etienne Migon (1643), where not only lines, but also planes are finally acting.

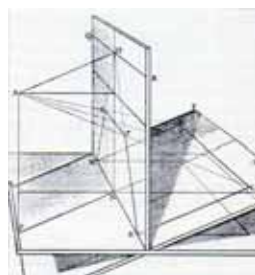


Figure 15: Etienne Migon, *La Perspective Speculative et Pratique* (1643) [16].

We will close the review of this period with another interesting practical treatise, written by Giulio Troili, or “Paradossi”, whose intriguing title is *Paradossi per praticare la prospettiva senza saperla (...)* appeared in 1672, literally “paradoxes to practice perspective without knowing it”. Contrary to the expectations, the treatise is very accurate and the constructions are clearly represented and explained [22].



Figure 16: The Paradossi's perspective setup including the distance circle (1672) [22].

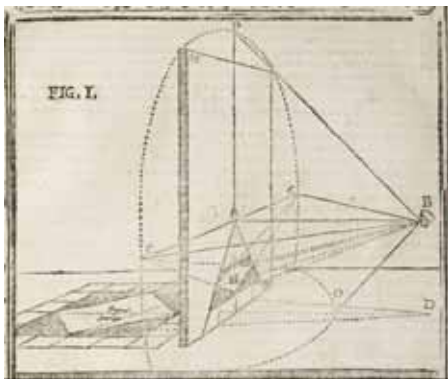


Figure 17: Another image of the Paradossi's perspective setup, showing the connection between picture plane and ground plane [22].

Remarkable in the two figures below (Fig.16 and Fig.17) is the distance circle, that is a circle drawn on the picture plane having the viewing distance as a radius, connecting the four distance points represented by Egnatio Danti into the Vignola's treatise (Fig.7). The picture plane setup shows now all the modern ingredients: ground line, horizon line, main point and distance circle.

The first evidence of the distance circle has already appeared into the unpublished treatise *Prospettiva pratica* by Ludovico Cardi, or "Cigoli" (1612). All these elements would have been fully used in the next centuries. At the moment, in spite of the substantial theoretical advancements, the main goal seems to remain the correct projective construction of the figures, therefore theory is developed and managed according to this purpose.

But it is just in this period that the passage from the *fictional* era, aiming to provide images from the space by perspective, to the *functional* era, aiming to carry on researches on the space by perspective, had his early start.

4. THE MATURE AGE: FINAL ALGORITHM

At the beginning of the XVIII century the connections between true shapes and their images were clearly understood. Before taking its name, the homology was actually 'at work', either in the *construction* of perspective images (Fig.18, above), or in the *reconstruction* of true shapes from perspective images (Fig.18, below).

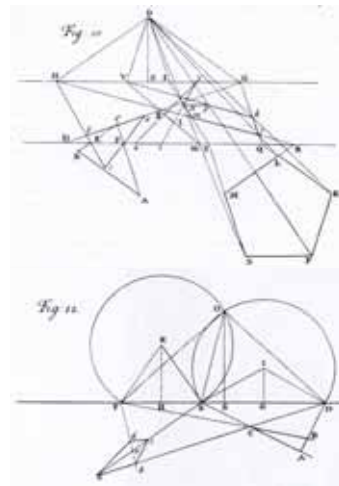


Figure 18: Brook Taylor (1749; 1st ed.1715). Perspective drawing, by given setup (above). Setup finding, by given perspective (below)[17].

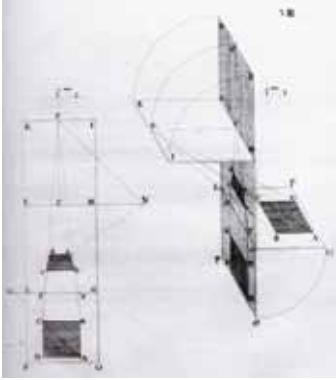


Figure 20: Joshua Kirby (1754). Image (left) vs space (right): folding viewing and subject planes into the picture plane [10].

The graphic procedures came from a clear understanding of the spatial setup of the geometrical elements and of their transformations (Fig.20), while new theoretical investigations were driving perspective towards its final ‘mutations’ as a branch of mathematics.

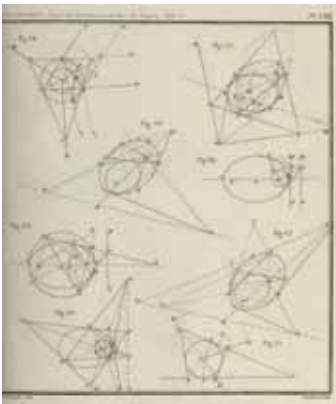


Figure 21: Jean Victor Poncelet (1865, 1st ed. 1822). Planche VIII [23].

Following the theoretical line starting with Euclid and empowered by Desargues, the *Traité des Propriétés projective des figures*, by

Jean Victor Poncelet, appeared in Paris in 1822, is the ‘baptismal’ work for a new geometrical field based on the principles of perspective, namely the Projective Geometry, a so widely general field that Harold Scott MacDonald Coxeter stated that even Descriptive Geometry, the new brilliant field codified by Gaspard Monge at the end of the XVIII century, has to be considered an *accident* of Projective Geometry. Getting rid of the ‘human’ visual restraints connected with perspective, the *projection/section* principle became a mathematical tool to fully ‘inflect’ in the abstract space, in search of new undiscovered properties. History testifies how important Projective Geometry would have been for the future advancements in Mathematics [1]. Focusing on the central projection, the projective invariants were finally defined as the bases of this new Geometry. Later on, thanks to Joseph Diez Gergonne (1847), also the fundamental principle of *duality* would have been found. With the Poncelet’s treatise, finally *homology* took its name and acquired its official role, not only in perspective, but also and more generally, in Geometry.

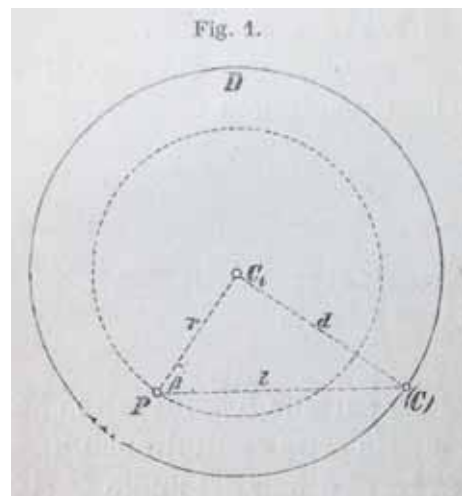


Figure 22: Wilhelm Fiedler (1874, 1st ed. 1871). Central projection, the graphic setup [8].

In few years Projective and Descriptive Geometry increasingly interacted, like in the brilliant work of Wilhelm Fiedler. Since in *Die Centralprojection als geometrisch*, published in 1860, he remarked that descriptive geometry should include more than the Monge's projections. In fact, the first figure in his masterpiece, the treatise *Die Darstellende Geometrie* appeared in 1871, shows the graphic setup of a central projection (fig.22), whose principle form the fundamentals of all the other projective methods. Starting from this figure, projector lines and angles, conflicting since the Euclid's theorem VIII, find a final appeasement, a clear definition and a reciprocal function. A special place is deserved to the *homology*, not only the one acting into the picture plane (supporting *drawing*), or the spatial homology between three-dimensional configurations (supporting *modeling*), them both available for scientific, as well for technical and artistic purposes, while in the last part of the book also the connections of this geometrical branch with the algebraic procedures through the projective coordinates is discussed. Last, the author does not forget to mention the aesthetic value of the projective methods, since he clearly states that by means of drawing (and models) descriptive (and projective) geometry combines *exactness* and *beauty*. The following four images show a wide field of possible homological applications, and the crucial role played by the graphic setup, that is, main point and distance circle.

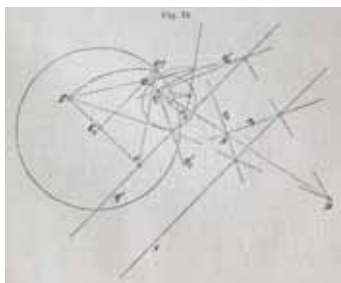


Figure 23: W. Fiedler. Homology: the graphic algorithm [8].

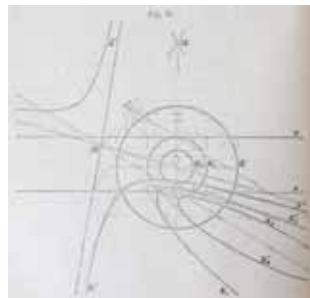


Figure 24: W. Fiedler. Homology in action in the abstract space [8].

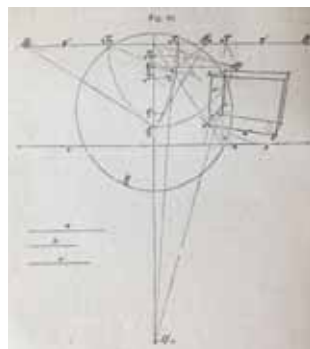


Figure 25: W. Fiedler. Homology in action in a perspective drawing [8].

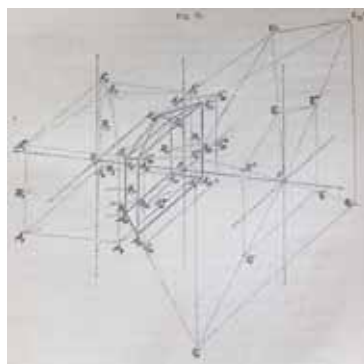


Figure 26: W. Fiedler. Homology in action in a three-dimensional perspective [8].

While mentioning these refined ‘visible’ images, we have to remind that few years before, thanks to the pivotal work *Geometrie der Lage* by Karl Georg Christian von Staudt (1847), Projective Geometry had been sublimated as a pure mental structure (*geometry of position*), independent on any metric contingency (*metrics geometry*) and on any graphic representation. Surprisingly, the lack of images works here as a powerful ‘amplifier’ for imagination, whose priority is to get the universal properties of the space before looking at the real configurations. That is also a recommendation he thoroughly address to the teachers, not to put students too soon in the middle of metrical details, but to help them to understanding what the space is and how it works. A recommendation resounding still valid nowadays.



Figure 27: Guido Hauck (1882). Figure and measure, final stage: the dominion of the space by means of the image [10].

We would like to complete our short excursus by mentioning the Guido Hauck’s treatise *Die malerische Perspektive*, published in Berlin in 1882, where the properties of the space appear completely dominated by a symphony of graphic constructions and virtuosic foreshortening. In this period also the applications had significant advancements, even in connection with photography, or the ‘automatic’ perspec-

tive of the XIX century, as the same Hauck’s work on the stereo-photogrammetric reconstructions from couples of photographs testifies. Of course, a new story to tell.

5. CONCLUSIONS

It is clear at this point the importance of the search of a secure geometrical connection between image and reality. This connection has been provided by the principle of *projection/section*, officially introduced in the early Renaissance, then enhanced thanks to Perspective, empowered by Descriptive Geometry, and finally widely generalized by Projective Geometry, that, taking into account *metrics* and *position*, also helped human beings to imagine new spaces and new geometries. The mutations occurred beneath the mutation of the perspective setup reveals this story, ‘dually’ developed between Art and Science, dancing between *fiction* and *function*. A story that should not be forgotten in the digital era, when new ‘dualities’ between *fiction* and *function* have emerged and are still emerging, that could provide to revitalize both Geometry and Graphics. All in all, into the screen is nowadays clearly recognizable the latest *pattern mutation* of the projective cast.

ACKNOWLEDGMENTS

The author would like to acknowledge the 16th ICGG committee and chairs.

REFERENCES

Books

- [1] C.B. Boyer. History of Mathematics. John Wiley & Sons, Hoboken, NJ, 1968.
- [2] F. Camerota. La prospettiva del Rinascimento. Arte, architettura, scienza. Mondadori Electa, Milano, 2006.
- [3] L. Cocchiarella. “Geometria Proiettiva”. Headword of the *Enciclopedia dell’ Architettura*, edited by G. Carbonara and G. Strappa, Wolters Kluwer Italia (to appear, 2014).
- [4] L. Cocchiarella. The Visual Language of Technique. Vol. 1-3. Springer, (to appear).

- [5] A. De Rosa, A. Sgrosso, A. Giordano. *La Geometria nell'Immagine. Storia dei metodi di rappresentazione*. Voll. 1-3. Utet, Torino, 2000-2002.
- [6] P. Della Francesca. *De Prospectiva Pingendi*. Edited by G. Nicco Fasola. Le Lettere, Firenze, 1984.
- [7] R. Evans. *The Projective Cast. Architecture and Its Three Geometries*. The MIT Press, Cambridge (Mass) and London, 1995.
- [8] W. Fiedler. *Die Darstellende geometrie in organischer verbindung mit der geometrie der lage*. Teubner, Leiptzig, 1871. Italian edition: A. Sayno, E. Padova (editors). *Trattato di Geometria Descrittiva*; del dr. Guglielmo Fiedler. Le Monnier, Firenze, 1874.
- [9] J.V. Field, J.J. Gray. *The geometrical work of Girard Desargues*. Springer, New York, 1987.
- [10] M. Kemp. *The Science of Art. Optical themes in western art from Brunelleschi to Seurat*. Yale University Press, New Haven and London, 1990.
- [11] G. E' Marchionibus Montis. *Perspectivae Libri Sex*. Apud Hieronymum Concordiam, Pisa, 1600. Italian edition: Rocco Sinisgalli (editor). *G. Del Monte. I Sei Libri Della Prospettiva di Guidobaldo Dei Marchesi Del Monte Dal Latino Tradotti Interpretati e Commentati da Rocco Sinisgalli. "L'Erma" di Bretschneider*, Roma, 1984.
- [12] E. Panofsky. *Die espektive als symbolische form*. Verlag Volker Spiess, Berlin, 1980.
- [13] R. Sinisgalli. *Verso una storia organica della prospettiva*. Edizioni Kappa, Roma, 2001.
- [14] K.G.C. von Staudt. *Geometrie Der Lage*. Korn, Nurnberg, 1847. Italian edition: M. Pieri (editor). *Geometria di posizione*; di Giorgio Carlo Cristiano von Staudt. Fratelli Bocca, Torino, 1889.
- [15] B. Taylor. *Linear Perspective: or, a new method representing justly all manner of objects as they appear to the eye in all situations*. Knaplock, London, 17815.
- [16] L. Vagnetti. *De naturali et artificiali perspectiva*. Bibliografia ragionata delle fonti teoriche e delle ricerche di storia della prospettiva; contributo alla formazione della conoscenza di un'idea razionale, nei suoi sviluppi da Euclide a Gaspard Monge. Libreria Editrice Fiorentina, Firenze, 1979.
- Web*
- [17] <http://www.aproged.pt/biblioteca/BrookTaylorlinearperspective.pdf>
- [18] <https://archive.org/details/dveregoledellap00vign>
- [19] <https://archive.org/stream/gvidivbaldiemarc00mont#page/52/mode/2up>
- [20] <http://bildsuche.digitale-sammlungen.de/index.html?c=viewer&bandnummer=bsb00007828&pimage=241&v=2p&nav=&l=en>
- [21] <http://digilib.netribe.it/bdr01/visore/index.php?pidCollection=De-prospectiva-pingendi:889&v=-1&pidObject=De-prospectiva-pingendi:889&page=001%20R>
- [22] <http://echo.mpiwg-berlin.mpg.de/ECHOdocu-View?mode=imagepath&url=/permanent/library/GKRQ3HY9/pageimg>
- [23] <https://ia600301.us.archive.org/20/items/traitdesproprio1poncuoft/traitdesproprio1poncuoft.pdf>

ABOUT THE AUTHOR

Architect, Phd, since 1999 he is Researcher (Assistant Professor) at the Politecnico di Milano, where he has been teaching at the Schools of Architettura e Società, and of Design. His research interests are in the field of Geometry and Graphics, their techniques and history including computer graphics, and graphic education. Member of the Department Board of the Department of Architecture and Urban Studies (DASTU) he can be reached by email at: luigi.cocchiarella@polimi.it.

PERSPECTIVE CONCEPTS. EXPLORING SEEING AND REPRESENTATION OF SPACE

Cornelie LEOPOLD

Technical University of Kaiserslautern, Germany

ABSTRACT: There had been a long tradition in analyzing the human seeing and conditions of perception in order to represent space and spatial objects similar to the way, we are seeing. Euclid tried already to describe the characteristics of perspective in his *Optics*. Alberti later in the Renaissance period referred to the ideas of Euclid and explained linear perspective as the section of the pyramid of vision. He described the image as practically received by the device "velum", the image plane, set between the eye and the object to be represented. There had been various seeing or perspective machines, which are able to represent the concept of perspective as a practical way to produce images of spatial objects according the process of seeing. Our students of architecture rebuilt some of these machines. The approach to perspective by the seeing machines explains the origin of the perspective concept from simulating the seeing process. The knowledge of optics and physiology of the eye had been part of the foundations of perspective. In this tradition can be analyzed the work of Guido Hauck, who tried to develop another concept of perspective, the "subjective perspective", which he based on the new physiological optics, received primarily by Hermann von Helmholtz. Following the geometric background of perspective we will observe the development of the comprehension of vanishing points. The clear concept of points of infinity and vanishing points initialized the development of projective geometry. Perspective is understood as a result of a transformation. Only then it was possible to develop the representation method relief perspective systematically, where half space is projected to a spatial layer. Jules de la Gournerie, Rudolf Staudigl and Ludwig Burnmeister worked it out systematically in the nineteenth century. We will ask and discuss, how the comprehension and concept of perspective had been developed in geometry and what had been the influences between theory and practice from various disciplines like geometry, art, architecture, physiology and perception theory.

Keywords: Perspective, Space, Representation, Seeing, Perspective Machine, Relief Perspective.

1. BASICS BY EUCLID AND ALBERTI

When we go back in the history of sciences, we meet the efforts to gain scientific knowledge and to understand the world around us. One instrument of our relation to the outer world is our seeing. Therefore it had been already an early question, how we see and what we can conclude from it.

Euclid developed a geometry of vision in his *Optics* [7] around 300 BC. The motivating force for his research had been the wish to derive statements about the distances of the planets and stars from the analyses of seeing,

therefore to gain scientific knowledge. In Euclid's *Optics* we can read for example: "Objects of equal size unequally distant appear unequal and the one lying nearer to the eye always appears larger. (...) Parallel lines, when seen from a distance, appear not to be equally distant from each other." [7, p.358]. These are already basic characteristics of perspective.

The reception of Euclid had been an important foundation in the Renaissance time. Alberti referred in *De Pictura* [1] to Euclid's *Elements*, when he started with the definitions of points and lines, and to Euclid's *Optics*, when he spoke about the characteristics of see-

ing, for example the changes of proportions in the image, and the angle of vision. Alberti wrote his *De Pictura* as a practical guide for artists, painters, although he described to pick up the basics from the Mathematician: "To make clear my exposition in writing this brief commentary on painting, I will take first from the mathematicians those things with which my subject is concerned. (...) In all this discussion, I beg you to consider me not as a mathematician but as a painter writing of these things. (...) The painter is concerned solely with representing what can be seen." [1, Book 1].

An important historical step had been, that Alberti explained clearly the practical creation of the perspective image by the section of the pyramid of vision with the help of the device "velum", the image plane, set between the eye and the object to be represented. "Nothing can be found, so I think, which is more useful than that veil which among my friends I call an intersection. (...) This veil I place between the eye and the thing seen, so the visual pyramid penetrates through the thinness of the veil." [1, Book 2]. It is passed on, that Alberti made impressive demonstrations of his indicated way to produce perspective images.

Whereas Brunelleschi verified the perspective paintings with his demonstrations by comparing them with the spectators view on the object, Alberti gave the initial point for developing methods to create a perspective image as a representation of a spatial object according seeing the object.

2. PERSPECTIVE MACHINES

To create images of spatial objects according a concept for the process of seeing leads us to several historical perspective machines, which bring again optics and the geometry of perspective close together. By analyzing these drawing and optical devices we grasp the origin of perspective as a representation of space from simulating the seeing process. We studied some of these perspective machines with our students of architecture to rebuild them for exploring their use during the night of sciences at our

university for a public audience in April 2014. The audience had been surprisingly highly interested to test and explore the machines. The students prepared models of El Lissitzky's "Wolkenbügel" ("cloud-irons"), 1923-25, as objects for the demonstration of the perspective machines.



Figure 1: Model of the "Wolkenbügel" during the science night

2.1 Brunelleschi's Mirror Device

There exist only descriptions of Brunelleschi's experiment [6], [13]. It had been told that Brunelleschi presented 1425 a perspective demonstration of the Florentine Baptistery. The aim of the demonstration had been, to show that his perspective paintings delivered the same image as looking at the real building from the specific viewpoint. A spectator had to stand in front of the Baptistery with his painting of the building on a panel with a small hole. The spectator had "to peer through this hole from the back of the panel at a mirror held in such a way as to reflect the painted surface." [13, p.13]. The device does not help to draw a perspective, but has the aim to verify the perspective painting or drawing and to convince the spectators of the accuracy of the perspective image when compared with the view to the real object. This mirror experiment only worked with an ax-symmetric building.

It is told that Brunelleschi worked for the demonstration of his "Palazzo de' Signori" without a mirror and the hole, but cut away the

area of the sky above the building in his drawing to enhance the illusion of the perspective drawing by merging image and reality.



Figure 2: Replication of Brunelleschi's Mirror Device

2.2 Dürer's Perspective Machines

Dürer's perspective machines however appear as active devices to draw perspectives. The historical researches came to the result, that Dürer experimented lately since 1514 with perspective machines. More details are presented in the paper about Dürer's contribution to perspective [15]. In his "Underweysung" Dürer described on the last pages four devices to draw perspectives, which have some similarities but also differences. The first two are already presented in the first edition of 1525, the other two additionally in the second edition of 1538 [5]. We picked up two to rebuild them for the science night. One of these machines can be understood as the materialization of the drawing method. The drawing method, which Dürer developed in two assigned views, was transferred to a spatial interpretation in the perspective machine. The string represents the straight drawn line. The original object gets scanned point by point with the help of the string and its intersection point with the image plane, like the "velum" of Alberti.



Figure 3: Replication of Dürer's perspective machine

The machine helps to understand the concept of perspective but is not really applicable as a drawing instrument, because it takes a long time to get the image point by point.



Figure 4: View through the grid of the rebuilt Dürer's drawing device

The students rebuilt and used a second well-known drawing device of Dürer, where a grid is placed in front of the object and the squared drawing surface corresponds with this grid, but has not to be in the same scale as the image frame. The eye position is fixed with the help of a stick or hole. The device has the aim supporting to draw what you see. The device is a drawing tool with the advantage of working

with different scales, so that the image can be drawn according to the seen object in a bigger or smaller scale. By using this device you do not feel comfortable to compare always the seen object through the grid with the grid on the drawing paper. The resulting wish to superpose the viewed and the drawn object is realized in the camera obscura and the camera lucida.

2.3 Camera Obscura

With the camera obscura we make again a step in the history of optics. It is an optical device, where the spatial surrounding is projected on an image medium. The device can be a box or a walk-in spatial installation where light is falling through a small hole on the opposite wall, the prototype of a photo camera. We built and used both, small pinhole cameras and a walk-in camera obscura in a workshop with Marek Pozniak, photographer and artist.



Figure 5: Photo, made by a small pinhole cardboard box camera

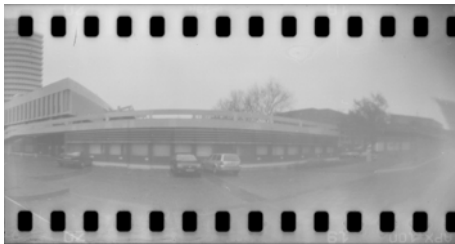


Figure 6: Panoramic photo, made by a small pinhole film can camera

It is surprising what we can get by a simple paper box with a pinhole (Figure 5). Instead of

a box we used also an old film can. The image then is projected on a cylinder and we receive a panoramic photo (Figure 6).

The camera obscura had been significantly improved in the history of photo cameras by integrating a lens in the hole of the camera obscura. Depending on the lens we receive deformations of the spatial objects. It had been interesting to realize that the image in the camera obscura on a transparent paper is upside down with the effect that you are forced to draw what you see and not what you think, because you do not understand the object in all details when seen upside down.



Figure 7: Drawing in the walk-in camera obscura

By using the camera obscura we comprehend the received image by the method of projection, by light, which gives us another fundamental concept of perspective.

2.4 Camera Lucida

The term camera lucida (lat. light chamber) is opposed to the term camera obscura (lat. dark chamber) with reference to the older device.

The English physician and chemist William Hyde Wollaston [13], [19] developed the first camera lucida in 1807. He used a prism with four optical faces in order to produce two successive reflections, so that the image is not inverted or reversed. The seen image is superimposed with the drawing surface. Our students prepared their own prism for a camera lucida.

Additionally, we used the NeoLucida by Pablo Garcia and Golan Levin [9] as well as a webcam version.

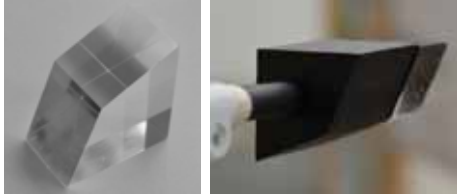


Figure 8: Prism of a self-made camera lucida



Figure 9: Drawing with the help of a NeoLucida [9]

2.5 Lambert's Perspectograph

With the Perspectograph of Johann Heinrich Lambert of 1752 we are focusing on the perspective transformation of a plane figure. It is the mechanical transfer of the relations between a plane figure and its perspective image figure. Lambert showed for example in his drawing (Figure 10) the perspective transformation of the ground plan of a garden. Collineation describes the relationship between the original figure and the image figure. With the Perspectograph we shift away from an artistic approach of rebuilding the seeing process to a mathematical understanding of perspective as a transformation. The basis can be found in Gi-

rard Desargues' famous theorem of 1639, fundament of projective geometry, where the relationship between two images of a figure is observed instead of the relationship between original and image [19].

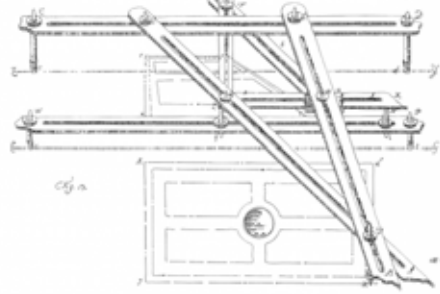


Figure 10: Lambert's drawing of the Perspectograph [14, p.161ff.]

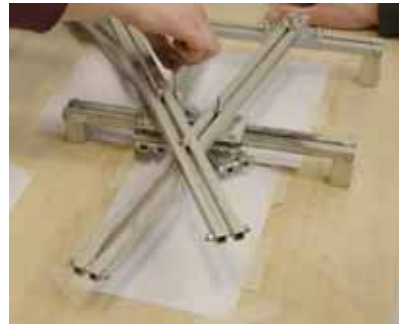


Figure 11: Replication of Lambert's Perspectograph

Precise explanation, high quality replication and digital animation of the Perspectograph can be found in the extensive material of the *Perspectiva Artificialis* project of University of Modena and Reggio Emilia [4].

3. SUBJECTIVE PERSPECTIVE

Erwin Panofsky criticised in his highly noticed article *Perspective as Symbolic Form* of 1924/25 [18] perspective theory as a rational theory, boldly abstracted from reality, far away

from an actual subjective visual impression with the assumptions that we see with one single unmoving eye and that our seen image would be adequately represented as the section of the pyramid of vision. The structure of an infinite, continuous and homogeneous mathematical space would be opposite to the structure of the psychophysiological space. He states: "In a sense, perspective transforms psychophysiological space into mathematical space. (...) It takes no account of the enormous difference between the psychologically conditioned 'visual image' through which the visible world is brought to our consciousness, and the mechanically conditioned 'retinal image', which paints itself upon our physical eye. (...) Finally, perspectival construction ignores the crucial circumstance that this retinal image – entirely apart from the fact that the eyes move – is a projection not on a flat but on a concave surface." [18, p.31]. But what could be the geometric solution?

Panofsky refers to the work of Guido Hauck of 1879 *Die subjektive Perspektive und die horizontalen Curvaturen des dorischen Stils* [11], where Hauck tries to develop a "subjective perspective", combining the mathematical and the aesthetic viewpoint based on the modern physiological optics. Hauck criticizes that the new achievements of the physiological optics did not effect the progression of perspective because of the apparent confirmation of the camera-obscura-images by the development of photography.

Hermann von Helmholtz provided as physiologist a new basis for the science of seeing. He dealt with all human types of sensation as basis for cognition. In refer to seeing he concluded in his *Handbuch der physiologischen Optik* [12] that although the retina receives the optical image like a camera obscura, the nerve cells, connected with the retina, effect seeing not the eye itself. Perceptions of external objects were seen as acts of our ability of imagination, as psychic activity. He emphasized that we learn through experiences, we perceive with various sensations and perceptions, we make

images of an object, if we move our eyes or body and view the object from various sides or touch, etc. the object. The perception of the object is the epitome of all these possible sensations. The perceived images are automatically connected with our imaginations and experiences. Thus an active instead of a passive viewer is assumed.

Hauck looks back in the history of the Renaissance perspective and analyses that many of the strong perspective construction rules were broken in the paintings, like in Raffael's School of Athens for example the representation the persons and objects at the sides, in order to achieve a satisfactory image. His method of subjective perspective directs towards a satisfactory image according the perception of the object. He sets the following conditions:

1. Principle of collinearity: Each line perceived as straight line should be also represented straight.
2. Principle of verticality: Perceived vertical lines should appear as verticals.
3. Principle of conformity: The apparent length of a line segment is proportional to the angle of vision.

But not all conditions could be completely fulfilled at the same time. Collinearity and conformity are contradictory. He demands the conformity only for the most important lines in the image and the collinearity only for the vertical lines and the horizon. Various perspective systems are his results. Perspective for him teaches the creation of compromises in the conflict of the condition of collinearity and conformity. The collinear-perspective system with conformity only around the main point of the perspective corresponds with the traditional central perspective. As a second system he suggests the conform-perspective system corresponding with the subjective image of perception.

The construction of such a subjective perspective is based on the idea of the development of the retina as part of a sphere. Hauck follows the idea of using cylindrical mapping of mathematical cartography. Georg Glaeser

calls those perspectives derived from a spherical image surface "transformed spherical perspectives" [10].

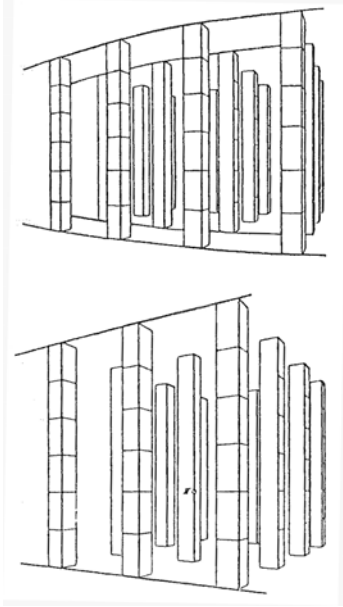


Figure 12: Piers hall in conform- (above) and collinear- (below) perspective [11], [18]

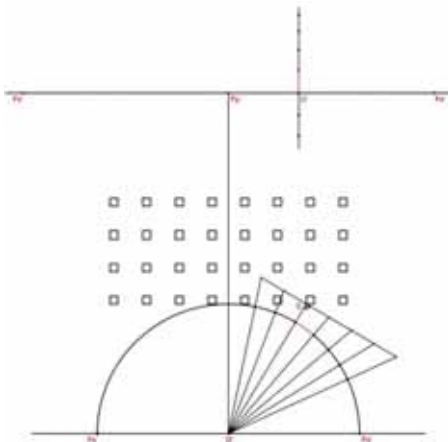


Figure 13: Construction method of the subjective perspective according to Hauck [11]

The semicircle represents the semispherical retina in top view and front view. The eye O is located in the centre. Hauck starts with the rectified semicircle as the horizon in the conform-perspective. The perspective image point is received by trace point method according traditional perspective but with the semisphere as image surface and with the difference that the determined distorted heights are transferred on a vertical through the image location on the horizon.

It was a remarkable attempt developing a new concept of perspective according the new psychophysiological knowledge, but the mixture of geometric method and arbitrary decisions remain unsatisfying. This may be also the reason that Hauck had attracted a particular attention by art experts but not by mathematicians. Geometry is not an empirical science. Henri Poincaré described this difference between geometrical and representative space in the triple form – visual, tactile, and motor - in his work *Science and Hypothesis* in a very clear way: "Representative space is only an image of geometrical space, an image deformed by a kind of perspective, and we can only represent to ourselves objects by making them obey the laws of this perspective." [20, p.57].

4. PERSPECTIVE TRANSFORMATION

For the further development of perspective in geometrical space the comprehension of vanishing points had been an important step. Guidobaldo del Monte [17] had been the first to introduce the term "punctum concursus". Guidobaldo characterized it as the point in which the images of parallel lines converge. But the most important step had been that he described to find the "punctum concursus" of the images of parallel lines with the help of the intersection point of a parallel line through the point of view with the image plane. Then the image of a line not parallel to the image plane can be constructed with the help of its trace point and vanishing point, although he did not use the term vanishing point. Using perspective representations for his drawings to explain the spa-

tial concept had complicated the comprehension of his explanations. In his drawing (Figure 14) point A is the point of view and X is received as the "punctum concursus" of the parallel lines BC, DE, FG by a parallel line through the view point A. The points B, D and F are the trace points of the lines.

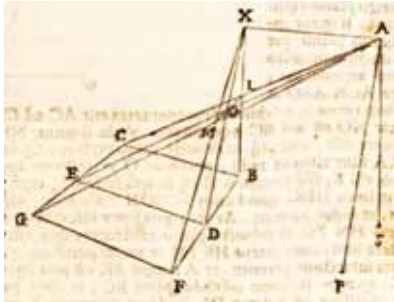


Figure 14: Perspective images of parallel lines through "punctum concursus" [17, p.42]

Brook Taylor, who handled the vanishing point in full general terms in his book *Linear perspective* of 1715, had done the next step. Although he characterized linear perspective as the art of describing exactly the representations of any given objects, as they would appear from any given point, he worked out a mathematical approach to perspective. He defined the most important used terms. The vanishing point had been defined as: "the point where the visual ray which is parallel to any original line cuts the plane of projection." [23, p.18]. And then he indicated all special cases and defined also a vanishing line. For his definitions and explanations he used the general case, where the plane of projection is oblique. Taylor wrote down his theory of perspective in a sequence of definitions, theorems and proofs. The most important step had been the general definition of vanishing points and lines. The concept of vanishing points and the deeper comprehension of the relation between the original figure and its perspective image had initialized the development of projective geometry.

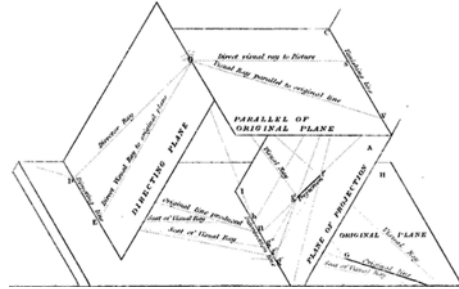


Figure 15: Definitions of perspective terms by Taylor [23, Fig.13, Plate I]

Lambert's studies in refer to his *Perspectograph*, the analyses of Jean Victor Poncelet in *Traité des propriétés projectives des figures*, 1822, and Jakob Steiner's thoughts in *Systematische Entwicklung der Abhängigkeit geometrischer Gestalten von einander*, 1832, lead finally to the system of projective geometry, as Karl Georg Christian von Staudt worked it out in *Geometrie der Lage*, 1847. He made a dedicated difference between a geometry of position and a geometry of measure. In refer to our topic the introduction of points and lines at infinity had been the most important step for the final clear understanding of vanishing points. Von Staudt wrote: "Two straight lines, lying in one plane, have either one common point or a common direction. Two different planes have either a common straight line or a common position." [22, p.23] (translated by C.L.). And he called his approach to geometry "perspektivisch" (perspectival).

This understanding gave finally the background for analyzing the projections themselves as transformations by studying geometric properties that are invariant under affine or projective transformations.

5. RELIEF PERSPECTIVE

Relief perspectives had been used for theater stages, for example "Teatro Olimpico" in Vicenza, 1585, or as built relief perspectives already in the 15th/16th century like "Chiesa di Santa Maria presso San Satiro" in Milan, 1479-99, but a systematic geometric approach

to relief perspective could be achieved only due to the comprehension of perspective as a transformation of spatial objects. Then the perspective transformation or collineation is applied to the spatial figure. The half space behind a front plane is transformed to the layer between the front plane and a parallel vanishing plane. The infinite half space is transformed in a finite space. The idea of the perspective transformation of a plane figure like in Lambert's Perspectograph is applied to a spatial object or the half space.

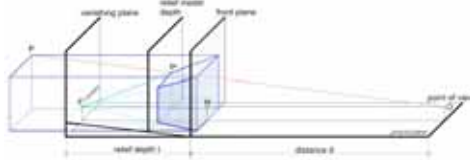


Figure 16: Concept of relief perspective transforming half space in a space layer [8]

The fundamental works for relief perspective *Traité de Perspective Linéaire* by Jules de la Gourmerie, *Grundzüge der Reliefperspektive* by Rudolf Staudigl [21], and *Grundzüge der Reliefperspektive nebst Anwendung zur Herstellung reliefperspektivischer Modelle* by Ludwig Burmester [3] had been all published in the 19th century. The amazing examples of relief perspective models by Burmester, rebuilt by Daniel Lordick with the help of a 3D printer [16], show the perspective transformation of typical solids.

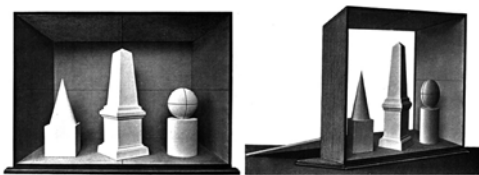


Figure 17: Typical solids in a relief perspective model by Burmester [3, plate IV, no. 1]

Staudigl pointed out the importance of relief perspective in refer to a systematic approach to perspective transformation: "Apart from the

values which have such studies for the sculptor, the same should also be of interest for those who devote themselves to the study of Descriptive Geometry, because the relief perspective is the most general method of projection, from which the orthogonal, the oblique and the perspective projection arise as special cases." [21, p.IIf.] (translated by C.L.). The relief depth, the space layer between front plane and vanishing plane, determines substantially the relief perspective. If the relief depth is zero, we get the usual perspective.



Figure 18: Perspective and relief perspective

We explored the creation of relief perspective models with our students of architecture in the so-called "All School Charrette", where students of all semesters work together in small groups in a one-week project, introduced and supported by all professors of our faculty. James Frazer Stirling's Clore Gallery, designed and built 1980-85 in London, had been the topic for the architectural analyses. Two perspectives of the entrance hall by Stirling formed the basis of the idea to build a relief perspective model according these perspective drawings. The two perspectives represent the entrance hall from two opposite view directions. Stirling's design approach by the two perspectives was supposed to get adequately represented by the relief perspective models. Examples of the students' works are presented in Figure 19-21. The relief perspective model represents space from a specific point of view. The spatial model itself does not remain an independent object; it becomes related to the spatial perspective transformation. Therefore we come back to our topic from the beginning, where we followed historical research efforts to explore seeing and find representations according our seeing procedure.

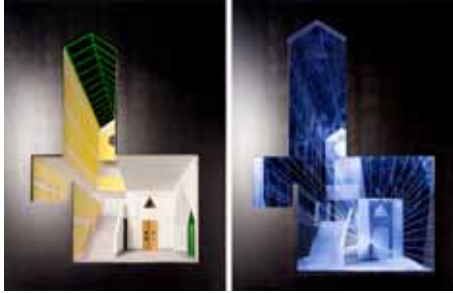


Figure 19: Students' relief perspective models (photos by Bernhard Frieze) [9]

When we look more in detail in Figure 20, we can comprehend the perspective deformation of the spatial objects, although it can be hardly shown in plane images.



Figure 20: Detail of the relief perspective model

Figure 19 and Figure 21 show the views in opposite directions in the entrance hall of the Clore Gallery. Additionally to the relief perspective models Figure 21 shows a corresponding perspective drawing. The relief perspective models had been built by using the same position of the viewpoint than in the perspective drawing.

The interest and stimulus in relief perspective can be found again in art, especially in applications for stage design. But only the geometric-mathematical development of vanishing elements, projective geometry and a transformational approach, systematically worked out in the 19th century, enabled a theoretical and applicable concept of relief perspective.



Figure 21: Perspective drawing and relief perspective model (photo by Bernhard Frieze)

6. CONCLUSIONS

By studying the perspective concepts and the representation of space according seeing, we get aware of the fruitful interrelations between theory and practice as well as the interactions between various disciplines like geometry, art, architecture, optics, physiology and perception theory. The optical and artistic approach to the perspective image entered a dialogue with the transformational and geometric-mathematical approach.

In the Renaissance time the epistemological conception of the world refers to the rediscovery and development of perspective theory. The philosopher Max Bense pointed to the relationship between art and philosophy: "Both, the perspective theory of the painter as well as the epistemology of the thinker, assume explicitly the subject-object relation thematically." [2, p.79] (translated by C.L.). And he concluded, that such a universal relation had been first established in the field of aesthetics and the philosophical treatment of the same problem lagged behind. Perspective theory and practice had been the motor for philosophical and epistemological questions.

It is interesting to see, that studying visual representation in art and science gave new impetus for our relationship to the world and the sciences as well as for the development of geometry in direction towards projective geometry as an important historical step in mathematics. Going back to these diverse roots of per-

spective in optics, art, geometry and mathematics remains an important foundation in our actual visual dominated culture.

ACKNOWLEDGEMENTS

The content of this paper had been worked out in refer to the topic of the seminar "I am a Camera" in winter semester 2013/14 for students of architecture, developed and realized together with Franziska Wilcken and Matthias Schirren, History and Theory of Architecture, TU Kaiserslautern. Marek Pozniak, photographer and artist, introduced us in the work with the different kinds of camera obscura. The All School Charrette "Stirling Hoch3" in summer semester 2013 gave me the chance to work on the topic relief perspective as the background for the students' models, according the initial idea of Bernd Meyerspeer, Building Construction and Design, TU Kaiserslautern. I herewith thank my colleagues for the stimulating exchange and collaboration as well as our students for their committed work.

REFERENCES

- [1] Leon Battista Alberti: *De Pictura*. 1435/36. Edition Latin/German, Darmstadt 2000. *On Painting*. (English)
www.noteaccess.com/Texts/Alberti/index.htm (06/2014).
- [2] Max Bense: *Konturen einer Geistesgeschichte der Mathematik*. II. *Die Mathematik in der Kunst*. Hamburg 1949.
- [3] Ludwig Burmester: *Grundzüge der Reliefperspektive nebst Anwendung zur Herstellung reliefperspektivischer Modelle*. Leipzig 1883.
- [4] Mara G. Bartolini Bussi et al.: *Perspectiva Artificialis*.
<http://archiviomacmat.unimore.it/PAWeb/Sito/Inglese/Templei.htm> (06/2014).
- [5] Albrecht Dürer: *Underweysung der Messung, mit dem Zirckel und Richtscheit, in Linien, Ebenen und gantzen corporen*. Nürnberg 1525, 2nd edition 1538.
- [6] Samuel Y. Edgerton: *The Renaissance Rediscovery of Linear Perspective*. New York 1975.
- [7] Euclid: *Optica*. The Optics of Euclid. Translated by Harry Edwin Burton. *Journal of the Optical Society of America*. Volume 35, Number 5, May 1943, 357–372.
- [8] Fachbereich Architektur (ed.): *All School Charrette 3. Stirling Hoch3*. Technische Universität Kaiserslautern 2013.
- [9] Pablo Garcia: *Machinedrawing - Drawingmachines*.
<http://pablogarcia.org/projects/machinedrawing-drawingmachines> (06/2014).
- [10] Georg Glaeser: *Extreme and Subjective Perspectives. Topics in Algebra, Analysis and Geometry*. BPR Médianácsadó BT/Budapest, 1999, 39–51.
- [11] Guido Hauck: *Die subjektive Perspektive und die horizontalen Curvaturen des dorischen Stils*. Stuttgart: Wittwer 1879.
- [12] Hermann von Helmholtz: *Handbuch der physiologischen Optik*. Leipzig 1867.
- [13] Martin Kemp: *The Science of Art. Optical Themes in Western Art from Brunelleschi to Seurat*. New Haven, London: Yale University 1990.
- [14] Johann Heinrich Lambert: *Schriften zur Perspektive*. Max Steck (ed.), Berlin 1943.
- [15] Cornelia Leopold: Albrecht Dürer's Contributions to the European Perspective Research Project in the Renaissance. In: Riccardo Migliari (ed.): *Prospettive Architettoniche*. Rom 2014 (forthcoming), 1–10.
- [16] Daniel Lordick: Reliefperspektivische Modelle aus dem 3D-Drucker. IBDG, Heft 1/2005. Innsbruck 2005, 33–42.
- [17] Guidobaldo del Monte: *Perspectivae libri sex*. Pesaro: Girolamo Concordia, 1600.
- [18] Erwin Panofsky: *Perspektive als symbolische Form* 1924/25, Leipzig Berlin 1927. English translation by Christopher S.

Wood: *Perspective as Symbolic Form*. New York 1991.

- [19] Alberto Perez-Gomez, Louise Pelletier: *Architectural Representation and the Perspective Hinge*. Cambridge, Mass.: MIT Press, 1997.
- [20] Henri Poincaré: *Science and Hypothesis*. London: W. Scott, 1905. Original: *La science et l'hypothèse*. Paris 1902.
- [21] Rudolf Staudigl: *Grundzüge der Reliefperspektive*. Wien 1868.
- [22] Karl Georg Christian von Staudt: *Geometrie der Lage*. Nürnberg 1847.
- [23] Brook Taylor: *Dr. Brook Taylor's principles of linear perspective, or, The art of designing upon a plane the representation of all sorts of objects: as they appear to the eye*. London 1835.

ABOUT THE AUTHOR

Cornelie Leopold studied Mathematics and Philosophy. She is academic director, head of the department Descriptive Geometry and Perspective in the Faculty of Architecture, Technical University of Kaiserslautern, Germany. Her research interests are: descriptive geometry, visualization of architecture, development of spatial visualization abilities, geometry and architectural design, structural thinking, philosophical background of architecture, geometry, and representation.

e-mail: cornelie.leopold@architektur.uni-kl.de
phone: +49-631-2052941 and postal address:
Technical University of Kaiserslautern, P.O.
Box 3049, D-67653 Kaiserslautern, Germany.
Website: www.architektur.uni-kl.de/geometrie

PERSPECTIVE: THEORIES AND EXPERIMENTS ON THE “VEDUTA VINCOLATA” (RESTRICTED SIGHT)

Riccardo MIGLIARI and Jessica ROMOR

'Sapienza' University of Rome, Italy

ABSTRACT: Erwin Panofsky's work on 'Perspective as symbolic form', has had a powerful impact not only on the art critics and therefore on the artistic-historical literature, but also on the studies that deal with the theme of perspective from the scientific point of view. The reflections stated by Panofsky in the incipit of his essay, in fact, were considered in an uncritical and superficial way to say that the perspective describes an image of the space that it represents, similar to that of the human vision only if: the eye of whoever is looking is positioned exactly in the projection centre used to generate the perspective image; the same eye remains motionless and therefore with the direction of the gaze perpendicular to the picture. This condition of observation of the perspective is known to the Italian scholars as '*veduta vincolata*'. Recent studies have proved, theoretically and experimentally, that, on the contrary: the eye of whoever is looking at a perspective can move in an area around a projection centre without causing a collapse of the perspective illusion, or better, of the sense of visual depth evoked by the perspective; the eye of whoever is looking can freely move around, in every possible direction, without compromising the effects of the perspective. These studies are expounded in this paper, first of all describing the theories and the experimentations that have given the above mentioned results. Secondly, it describes the verifications carried out on important perspectives painted on walls and entire rooms, in which the conditions of one single projection centre are respected. It also describes other works, in which appropriate solutions permit to dilate the area of the *veduta vincolata*. Finally, in a quick re-reading of the first pages of Panofsky's essay, this paper aims at an interpretation of it, which attempt to overcome the conflict between the advocates and the detractors of perspective as legitimate form.

Keywords: Perspective, Panofsky, Restricted Sight, Veduta Vincolata.

1. PANOFSKY AND THE RESTRICTED SIGHT

Erwin Panofsky's work on 'Perspective as symbolic form', has had a powerful impact not only on the art critics, and consequently on the artistic-historical literature, but also on the studies that deal with the theme of perspective from the scientific point of view.

The reflections stated by Panofsky in the incipit of his essay were, in fact, at times, assumed in an uncritical and superficial way, asserting that perspective only describes an image of the space that it represents, similar to that of the human vision, if:

- the eye of whoever is looking is positioned exactly in the projection centre that is used

to generate the perspective image;

- the same eye remains motionless and therefore with the direction of the gaze perpendicular to the picture plane.

The observation condition of the perspective, in which the eye is placed exactly in the projection centre, though not necessarily motionless, but free to rotate, is known to the Italian scholars as '*veduta vincolata*' (restricted sight).

The aforesaid interpretation of Panofsky's work, combined with the high esteem that its Author deservedly is held in, has had weighty consequences, like those concerning a strictly axial and static reading of the architecture of the Renaissance period, along with a decline of perspective, as a planning verification tool, that has

hit all the architecture of the modern¹ movement. Perspective appeared to be a misleading tool, culpable of misshaping the represented objects and unable to simulate the human perception of the space.

This prejudice towards perspective appears, today, to be totally anachronistic, particularly if we consider the continual use that our contemporaneous world makes of the perspective image in every bi-dimensional expression of space and in all its applications, both ludic and professional².

2. FOR A NEW READING OF PANOFSKY'S TEXT

We would therefore, first of all, like to suggest a different way of reading the part of Panofsky's text that is at the root of the above-mentioned misinterpretation.

In fact, in the passage where he mentions the characteristics of the 'Zentralperspektive', recalling the constraint of the sight, Panofsky explicitly refers to the two earliest centuries of perspective's history, namely the period from the Renaissance to Desargues³, and he uses expressions that clearly are related to the aforesaid restrictions to the rules established at that time (particularly in the proposition XXX of *De Prospectiva Pingendi*). In other words, the passage in question, if carefully re-read, tells us that Panofsky, being a son of his time, did not agree with this idea of the founding fathers of the legitimate construction.

But why does Panofsky confine this conception of perspective to Desargues and not, earlier, for instance, to Guidobaldo del Monte, assuming that this last found the construction of the vanishing point? Because, while Guidobaldo only found a geometrical rule that mechanically justifies the convergence of the images of parallel straight-lines, Desargues defines the points and the straight-lines at infinity, as a support of



Figure 1- The vault of the Sant'Ignazio Church in Rome, frescoed by Andrea Pozzo, is one of the several 'Glories', a pictorial genre that were very common during the 17th and 18th century; in this fresco, thanks to the linear perspective, to the surface and the reduction of the apparent sizes of the figures, skilfully controlled, is induced the sense of ascending into the depth of the sky.

classes of parallel straight-lines and planes, and thus explains what the vanishing points represent.

With this attainment, perspective becomes a tool capable to deal with infinity in finite terms. An epochal change took therefore place at that time, modifying profoundly on the one hand the conception of space and on the other the human ability to represent it. To experience this different conception and expressive capability, it is enough to visit one of the Roman Baroque churches and to turn one's eyes towards the vaults, where the figures precipitate attracted into the depths of the sky by an inversion of the force of gravity (fig. 1).

But let us get back to Panofsky and his 'restricted sight'.

Another limit of the legitimate construction, or better of the Renaissance perspective, is that of its relation to space, considered as an isotropic and homogenous continuum. Whereas on the contrary, admonish Panofsky, the perceptive space is anisotropic and not homogeneous.

To understand this idea, we imagined two different situations:

In the first there is the Cartesian space, in which is immersed a plane that becomes support

¹ See, in particular, the essays written by Alberto Sartoris and Bruno Zevi.

² We refer to the architectural rendering, but also to more elaborated applications which go from the introduction of virtual models in the film shooting (VFX), to the photogrammetric survey of last generation (IBM, Image Based

Modelling).

³ "This correct construction was in fact invented in the Renaissance, and although later subjected to various technical improvements and simplifications, it nevertheless remained in its premises and goals unchanged to the time of Desargues."

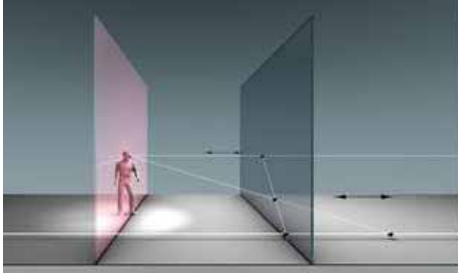


Figure 2 - In the Renaissance conception, perspective is the intersection with a plane of the visual pyramid; the objects to be represented (r,P), the observer (O), the picture plane (π') and the perspective itself ($r'P'$) are immersed in a Cartesian space, homogeneous and isotropic.

to the perspective, generated by sectioning the visual pyramid: the perspective is thus a two-dimensional geometrical structure (fig. 2).

In the second there are two spaces overlapping in projective correspondence, one is isotropic, the other anisotropic (solid perspective)⁴ (fig. 3).

Well then: the Renaissance perspective identifies itself in the first case and, particularly, in the second of the procedures proposed by Piero della Francesca, as well as by many other Authors: the one that uses the orthogonal projections to generate a plane perspective; this procedure operates in the isotropic space and generates, therefore, a plane and static image;

In the second case, instead, it identifies itself in the solid perspective of Borromini and later in that still solid but theoretical of Wilhelm Fiedler, who, thanks to the projective correspondences which come from the concept of infinity, transforms the isotropic space of the Renaissance conception, into the anisotropic space of Baroque conception, in all its dynamism.

With Panofsky's words, we could say that

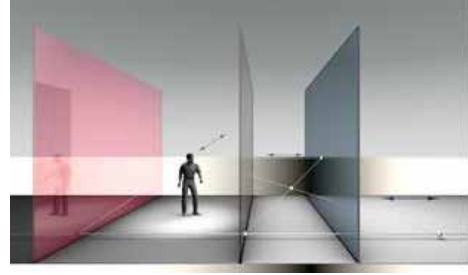


Figure 3 - In the opinion of Panofsky the psycho-physiological space of the visual perception is unhomogeneous and anisotropic, like the space of the solid perspective is.

the solid perspective transforms the mathematical space into a psycho-physiological space, and vice versa.

It is finally interesting to note that, if Panofsky is able to measure the limits of the Renaissance conception, it is because Desargues enabled him to do it, as well as he enables all of us today. Thanks to the creation of a geometry capable to include infinity among its axioms.

Now let us get back to the Renaissance perspective eye, which is fixed and motionless. Panofsky observes that that perspective is unable to reproduce the phenomenon of the vision, exactly because it does not consider the mobility of the eye (besides of the binocular vision). Well, this is the perspective, the one in between Alberti and Guidobaldo, which does not consider the mobility of the eye or, to be clearer, which is unable to consider it.

At this point of the essay it is reasonable to expect a surmounting, an answer to the question that arises in the reader, and namely: if not this, then, which perspective?

Here comes into play the mobility of the eye, because the free rotation of the eye, around a

⁴ The solid perspective, which is the basis of the modern conception of perspective, puts into perspective relation two overlapping spaces: one, the object of the representation, is homogeneous and isotropic, the other which supports the representation is contracted and therefore unhomogeneous and anisotropic. The observer (O), the limit plane (α), the objects to be represented (r,P), are immersed in the isotropic space of the reality, which here is

represented by an even background; the vanishing plane (π') and the perspectives of the represented objects (r',P') are immersed in the anisotropic space, here represented by a veiled background that alludes to its contraction. The plane of the traces ($\tau=\tau'$), locus of united points and straight-lines, belongs to both spaces.



Figure 4 - In the rotation of the eye, which pivots in the projection centre to look at the surrounding space, the visual field spontaneously assumes a spheroidal appearance, due to the convergence of the straight-lines that are parallel to the picture plane.

projection centre, 'gives to the field of vision a spheroid shape'. And up to this point we can only agree with all the Panofskyan reasoning. In fact, if we observe a plane perspective from its projection centre, looking freely towards the right and the left hand side, and up and down, we will see that the straight-lines that at first appeared to be parallel and horizontal, being such with respect to the picture plane, appear to converge now at the right, now at the left; and those which appeared to be parallel and vertical, appear to converge now at the top, now at the bottom; and that, in the continuity of this transformation, the visual field seems to assume a spheroidal appearance (fig. 4).

As is well-known, Panofsky uses considerations of a physiological nature to explain this phenomenon, but it can all easily be explained even within the rules of perspective as we know it today, and following a path that seems to be not only simpler, but able to reconcile the legitimate Renaissance construction to the Baroque construction and the antique perspective to the modern one, extended to the projective space. Now we come to the crucial point that, in our opinion, constitutes the only misinterpretation of which Panofsky was a victim.

The points that stimulated his attention were two: the marginal aberrations and the curvature of the retina, which would cause the curvature of the visual space. A closer observation of these two reasoning may perhaps take us too far from our objectives and nevertheless, very briefly, it

is necessary to mention the following.

As for what concerns the marginal aberrations: these will be seen only if the perspective is observed from a viewpoint different and very far from the projection centre. The marginal aberrations belong therefore to the rational and isotropic space, and not to the anisotropic and non-homogeneous space typical of the perception. And it is not just chance that this phenomenon can skilfully be used in the anamorphosis, exactly thanks to a displacement of the observation point, with respect to the projection centre.

Panofsky grounds his theories on the eighth proposition of Euclid's *Optics*, where Euclid affirms that straight-line segments, parallel to each other and placed at different distances from the eye, are seen under angles which are not proportional to the respective distances from the same eye. This means, for instance, that if the segment **AB** is situated at a distance from the eye which is twice the distance of segment **CD**, the angle subtended by **CD** in the eye is not twice the angle subtended by **AB**. Beyond the debate that has heated up about this proposition and its historical significance⁵, it is of interest to observe that this geometrical relation precedes the perspective, understood as section of the visual pyramid. Using a modern language, which is more familiar to us, we would say that the proposition describes the projection operation, but not yet the section operation. An when we carry out the section operation by means of any surface, leaving out the sphere centred in the eye, then the projections of the aforesaid segments will have lengths which do not represent the same ratios expressed by the angles subtended by them. All this, however, in an evaluation that is abstract and unrelated to the perspective. Because, if we restore the restricted sight, then those straight-line segments, whichever is the surface to which they belong, subtend in the eye the angles that Euclid talks about.

As for the curvature of the retina, the studies of the transactionalist school (Ames and others) have shown that the shape of the retina does not

⁵ We refer, in particular, to the studies carried out by C.D. Brownson (1981) on the compatibility between Euclid's

Optics and the linear perspective.



Figure 5 - Two images of the original experiment carried out by Ames.

have any importance as far as the visual process is concerned; we refer, in particular, to the experiments conducted using aniseikonic glasses, but also to the famous distorted room (known as 'the Ames room') (fig. 5), to the rotating trapezoid and many others. As it is determined in a theory that stably endures since the 1950s, there exists a world, made of solid shapes and referable to geometrical models, which belongs to the experience of all the senses and of the reason. Our brain compares this Cartesian world with the visual experience, which is subjective, in order to choose the formal models that best respond to the same experience. Therefore, when we observe the Ames room, we perceive a space that has a parallelepiped shape, because this is our everyday and most common experience, even if the room is of a pseudo-pyramidal (fig. 6) form. It is not just chance: it is exactly the perspective that shapes the room in a skilful play of false correspondences, misleading the eye and the perception of the space.

We will focus further on this ambiguity, because it is present in the architectural perspectives, which are the main topic of this study. In fact, as in the experience of Ames a simulated and misleading architecture (the room of regular shape) prevail on the perspective (that is, on the real shape of that room), so in the large *trompe l'oeil*, from Renaissance to Baroque, there are always two interpretive keys: the geometrical key and the architectural key.

These two keys can be coherent or not, and it can therefore happen that one perception strengthens the other, or that one (and it is always the architecture) corrects the other. As we will see shortly, these two manners have, in the illusionistic perspective, their champions: on the

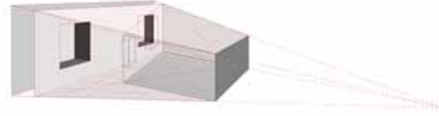


Figure 6 - The true shape of the Ames room and its geometrical construction.

one hand Andrea Pozzo, who uses perspective in full accordance with its geometrical code, on the other Agostino Tassi, who on the contrary does not respect that code, he just imitates it, deliberately introducing numerous errors or corrections, which nevertheless are not considerable because they are disguised by the dominant interpretative key, namely that of the architectural space.

In particular, to better understand, on the one hand Pozzo respects the uniqueness of the vanishing point of straight-lines that are image of parallel straight-lines in the illusory space, on the other Tassi skilfully bestows to those straight-lines a large number of vanishing points, as many as is needed to follow the onlooker in his movements inside the frescoed room. Therefore, on the one hand Pozzo takes advantage of the eye's rotation, pivoting in the projection centre and resorting to other expedients, mitigating the effects of a displacement of the observer; on the other hand Tassi disregards the code in order to nullify the effects of the displacement.

The perspectives painted by Tassi, thus, behave like the Ames room, because they appeal to the viewer's experience and to his mental models, in order to hide the derogations imposed to the perspective rules and the true shape of the space that these perspectives describe, if interpreted literally.

All this to say that, if we want to go more thoroughly into these aspects of perspective and its applications to the art of the *trompe l'oeil*, we do need to reason about the question that we posed at the beginning of this paper, and namely on what the 'restricted sight' is and on what its limits are; on how, in certain cases, these limits can be overcome and on how the related effects can be experimentally reproduced and measured.

We have thus created an interactive model, which permits us to reproduce in laboratory a condition that otherwise would have required much more effort. The model represents a bare



Figure 7 - The model used in the restricted sight experiment, as it really is. The wall that cuts it at a fourth of its length, houses an illusory perspective that represents the part of the colonnade that is hidden.

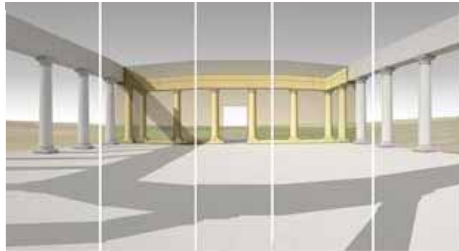


Figure 8 - The perspective, if observed from the projection centre, rotating the gaze freely around, simulates the sensation of curvature of the visual field, described by Panofsky.

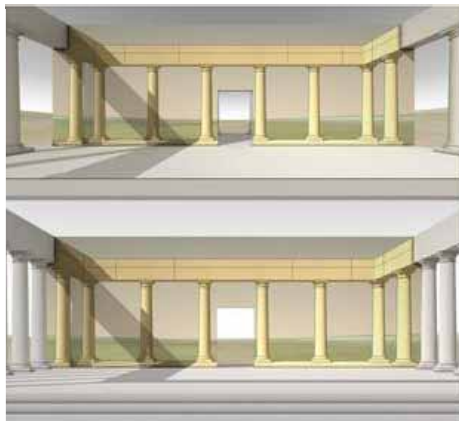


Figure 9 - The displacement of the viewer along the normal to the picture plane produces a contraction or a dilatation of the illusory space.

natural environment in which a peristyle that is formed of twenty-four Tuscan-Order (fig. 7) columns is placed. The peristyle is intersected by a wall, so that eight of the columns are located beyond it. On the wall is painted the perspective of the part of the peristyle that the wall is hiding. The onlooker can freely move about inside the scene, and should start standing exactly on the projection centre point.

The following experiences are possible:

- The rotation of the eyes of the viewer: the perspective perfectly simulates the illusory space in every condition, making it appear to be the natural prosecution of the real space; in particular, rotating the eyes to the left and to the right, up and down, we perceive the convergence of the images of the lines that were parallel in the normal observation, with a 'motionless eye', as Panofsky would have said, and we can see how, in the continuity of the vision, this mutable convergence gives a spheroidal shape to the visual field (fig. 8);
- The displacement of the viewer along the normal to the picture plane: the perspective continues to simulate a space that is perfectly contiguous to the real one, but it expands it or contracts it depending on whether the viewer is moving away from or nearer to the painting (fig. 9);
- The extent of the effects of the displacement of the viewer is described by a linear proportion (fig. 10): if the viewer is moving away from the picture plane twice the primary distance, the space simulated by the perspective is twice deeper than the real one; if the viewer halves his distance from the picture plane, the perspective contracts, by half, the

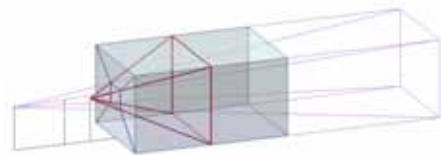


Figure 10 - The contraction or the dilatation of the illusory space are proportional to the displacement of the viewer, as this scheme easily shows us.



Figure 11 - The collimation between the illusory space and the real space is lost when the onlooker translates parallel to the picture plane: this is the phenomenon, together with the marginal aberrations, which induces Agostino Tassi to increase the number of vanishing points of the straight-lines that are perpendicular to the picture plane.

depth of the simulated space; if the viewer moves so close to the wall whereon the perspective is frescoed that he goes beyond the door opening that is located at the centre of the wall, then the perspective will collapse into the two dimensions of the plane;

- Finally, the displacement of the viewer along the parallel to the picture plane: the perspective rapidly loses the continuity with the simulated space and this is why the quadraturist painters increase the number of vanishing points of the straight-lines that are perpendicular to the picture plane, obtaining in this way an extension of the limits of the restricted sight (fig. 11). The multiplication of the above-mentioned vanishing points allows, moreover, to restrain the apparent deformations (the ones Panofsky calls 'marginal aberrations').

3. SOME APPLICATIONS: RESTRICTED SIGHT AND ARCHITECTURAL PERSPECTIVES

That said, we would like to apply the observations made so far to the investigation and analysis of the two emblematic cases that we have mentioned. The first is a perspective, situated in the Casa Professa del Gesù in Rome, painted by Andrea Pozzo during the years 1681 - 1686; the second is a room in Palazzo Lancellotti, also in Rome, frescoed by Agostino Tassi during the years 1617 - 1621. In both cases we deal with unitary perspectives, that is to say, they have in



Figure 12 - Panorama of the corridor that leads to Sant'Ignazio's rooms in the Casa Professa del Gesù, in Rome.

common the intention of simulating a single architectural space, even if this is painted onto various surfaces; as if the aforesaid illusory space were projected from the centre onto the whole host enveloping environment. This is why it would be helpful to coin a new term for this kind of perspectives, for instance poly-perspectives, or multi-faceted perspectives, considering the three-dimensional form of the support surface. The only difference between these two perspectives is in the freedom of the onlooker who, in the first case, freely can move his eyes in all directions, paying, though, the effects of his displacement, as we will see in a while; whereas in the second case the viewer not only is free to look in all directions, but also to move about in the room.

The corridor in the Casa Professa del Gesù is a long and narrow space, totally devoid of decorations in relief, and with a low-profile barrel vault ceiling. The end wall is oblique, with respect to the axis of the corridor. The viewer is standing, approximately, at the centre of the corridor and he therefore sees the two long side walls and the barrel vault strongly foreshortened (fig. 12).

The study of this multi-faceted perspective, unitary as we said, have to be done analyzing, independently one from another, the parts projected on the different enveloping surfaces, since each of these parts of the perspective has its own characteristics.

In the side walls (fig. 13), the perspective goes much beyond the limits set by the application of the art of perspective to the visual field



Figure 13 - View of a side wall along the normal to the picture plane: to take the whole painting in at a glance, it is necessary to dilate the visual angle until it measures 150 degrees.



Figure 14 - View of the same wall, with the direction of the gaze parallel to the picture plane.

of the viewer, limits of which we can find explicit treatment already in Piero della Francesca and which has then been handed down from treatise to treatise until the 20th century. As is known, some scholars attribute to this limit angle an aperture of 60 degrees, others attribute different values, each time explaining these statements in the most various ways (and the most inconsistent). Actually, the limit gains sense only in the case of the legitimate construction (from Renaissance to Desargues, as Panofsky recalls) because it makes sense to talk about

it only in the case in which the eye is obliged to look at the perspective, keeping the gaze perpendicular to the picture plane. And in this case too, we should have to make many explanations, which we have not space for here. But Andrea Pozzo very well knows what the constraints of the restricted sight are, and what not, so he takes it for granted that the viewer can look at his perspective with strong angulations and even with the direction of the gaze rotated at an angle of 90 degrees with respect to the walls and therefore



Figure 15 - View of the wall with the entrance that gives access to the corridor.

parallel to these (fig. 14). This leads to the presence of considerable apparent deformations in the most peripheral areas, deformations that, however, we call 'apparent' because they can only be seen if the onlooker moves away from the restricted-sight position. The 'lesson' Pozzo teaches us also helps to show, were it needed, that perspective and anamorphosis is the same thing: both obey the same rules.

The entrance wall (fig. 15) houses the only part of this perspective that does not have particular characteristics, unlike the end wall which is, as already mentioned, strongly oblique (fig. 16). Pozzo here paints an illusory space that widens and regularises the corridor, insomuch that the viewer gets the feeling of being standing in front of a perspective that only has one vanishing

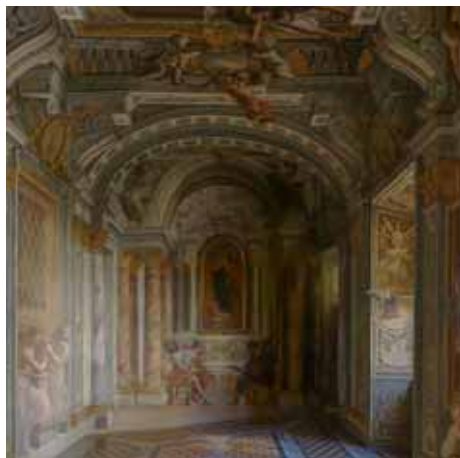


Figure 16 - View of the end wall, with the direction of the gaze parallel to the axis of the corridor, and therefore oblique with respect to the same wall.

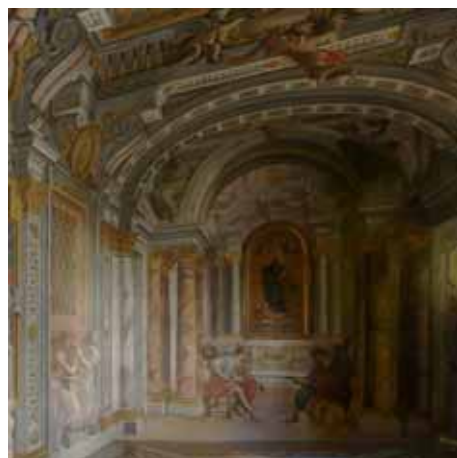


Figure 17 - Frontal view of the end wall, where the presence of two vanishing points is evident. This image is taken from the projection centre.

point, and precisely that of the images of the parallel straight-lines of the side walls of the corridor, which ideally proceed into the depth of the scene. Actually the perspective, being painted on a support that is oblique with respect to the sides of the corridor, uses two vanishing points (fig. 17).

Here we deal with another 'lesson' of Andrea Pozzo who shows us how conventional the classification of the perspectives in frontal, at an angle and of an inclined picture plane is, since a



Figure 18 - View of the vault; the perspective simulates corbels and rectilinear beams which evenly divide the spaces.

perspective of one kind transforms itself into the other when it is observed by freely rotating eyes: the experience of the colonnade very well emphasized this.

Finally, the barrel vault. Here the illusory effect reaches its highest level, because the perspective frescoed by Pozzo goes as far as to destroy and dematerialise the support, substituting the smooth surface of plaster with a rich framework of corbels and beams perfectly rectilinear, which evenly divide the deepest spaces in which angelic figures are moving (fig. 18).

How has this miracle of geometry been achieved? Very simply intersecting the vault surface with the planes that project the edges of the beams, as immediately will understand the viewer who has severed himself from the constraints of the sight.

Now, without lingering further on the geometric peculiarities of this work, and keeping to the point, we will consider the solutions utilized by Pozzo for the decorations of the corridor, to make them usable even in their most common use, that is, the visitors crossing of the corridor, which therefore is subject to a continuous and strong displacement of the viewer.

Among the lacunars of the ceiling there are angels that hold some medallions, on which are painted several portraits

This detail encourages a reading, bi-dimensional too, of the illusory space, almost as if it were the representation of a portrait gallery. And indeed, the sides of the corridor are also hosting some sacred representations that are in two dimensions and that therefore can be read in two ways:

- Standing on the projection centre these representations look as if they were paintings hanging in the illusory space of the gallery that the perspective represents;
- Moving about in the corridor, the illusion effect collapses, but not the paintings, which can be seen one by one as if they were part of the real space.

In this way the dynamic passage of the corridor retains an interest, even if the perception of depth vanishes, and the perspective, no longer



Figure 19 - Detail of the anamorphosis on the vault, seen standing on a spot that is different from the restricted sight position. readable, becomes a decoration of abstract beauty (fig. 19).

Different, but not less efficacious, the expedient used by Agostino Tassi at Palazzo Lancellotti (fig. 20). Here the artist does not want and does not allow the perspective to lose its effectiveness because of the onlooker's movement inside the space of the corridor. The perspective therefore follows the viewer in his movements

displaying the same number of images, skilfully merging them together the one with the other. The study of the fresco shows, in fact, a multiplication of the main point, or better, of the point which is the foot of the perpendicular drawn from the eye to the picture plane, and that coincides with the vanishing point of the straight-lines that are perpendicular to the picture plane. The effect that arises from this is amazing, but it can fully be understood only in a direct inspection. In fact, the dimensions of the corridor (fourteen by nine metres and nine metres and a half in height) are such that it is impossible to appreciate the perspective as a whole, looking at it from a single point of observation. The viewer, therefore, is induced to linger on the detail that is nearest to him and there he finds, unfailingly, a correct perspective; whereas the farther parts, which are not as much present from that observation point, perfectly hide the derogations that that point of observation, and only that one, opposes to the perspective rules.



Figure 20 - One of the walls of the room, situated in Palazzo Lancellotti, frescoed by Agostino Tassi.

4. CONCLUSIONS

To recapitulate, we have defined the following:

1. The modern perspective (with this term we allude to the perspective from Desargues onwards, to remain at Panofsky's scheme) admits the free rotation of the eye situated in the projection centre (restricted sight) and with this, and with the motion of the eye, it is capable of simulating even the impression of curvature of the visual field that Panofsky erroneously ascribes to the curvature of the retina.
2. We can ascribe the aforesaid error to the fact that Panofsky writes in 1927 and thus a quarter of century before the scientific achievements of the transactionalist school. Furthermore he is influenced by the discovery of the optical corrections adopted in the classical architecture and he believes that the ancients did apply the same corrections, with the same effects, to the perspective. In this supposed analogy, nevertheless, he does not consider the different sizes of the two artefacts, nor does he consider the visual fruition



Figure 21 - Verification of the cross-ratio in one of the perspectives of Second Style situated in the House of Augustus.

of the ones and of the others.

3. The restricted sight admits an ample displacement of the observation point in the normal direction to the picture plane, whereas it is much less tolerant for a displacement parallel to the picture plane, that emphasizes the 'marginal aberrations' and not only. This limit has been passed, during the 17th century, by artists like Agostino Tassi through the repetition of the primary point (which is also the vanishing point of the normals to the picture plane). The perspective, thus revised, transforms itself from one single perspective geometrically coherent into a whole of perspectives, each of which is coherent in itself, merged into one, like the single frames of the tracking shot in a cinematographic filming. This produces a multiplicity of vanishing points, like those observed by Panofsky in some of the Roman paintings of Second Style. It is interesting to note that the straight-lines that Panofsky interrupts on the vertical axis of the Roman perspective construction can instead be interrupted on the horizon (fig. 21), which is in line with what we have said above. Furthermore, Panofsky limits his verification efforts to the research of the vanishing points, whilst the research of a rule, in perspective, should always be extended to the reduction of the apparent sizes. Now, if the perspective is correct, from a projective point of view (legitimate, as Panofsky would have said), the cross-ratio of four points which define, in the illusory space, three equal intervals, has to be equal to four-thirds. Well, this measurement, made on many Pompeian and Roman (Casa di Augusto) perspectives, returns exactly this value, showing that, at least as far as the depths and their illusionistic rendering is concerned, the perspective is the one that the Renaissance artists have re-discovered, perhaps with different procedures, but with the same result.
4. Finally, the study conducted by Panofsky does not consider the two possible interpretive keys of an architectural perspective: on

the one hand the geometrical key, which reveals itself by means of an inverse procedure, capable of describing the shapes that are represented in space; on the other the architectural key, which obtains the same result simply observing the typical characteristics of an architecture, like the symmetry, the horizontality of the architraves, the verticality of the pillars, the proportions of the Order. When we observe such a perspective we receive a strong impression of depth because of the sum of these two effects and because of the motives that the Ames room experiment has very well emphasized. Now, these two keys can be coherent, that is, aiming at achieving the same result (like in Pozzo), or they can be independent of each other (like in Tassi), in which case the second prevails over the first because it appeals to a psychological mechanism, whereas the first can be verified only by means of reason. From this derives the great liberty that perspective gives the artists and all those who want to use perspective to simulate the space, 'bursting through the depths' of the wall surfaces, extending the architectural frameworks, projecting human figures into the depths of the sky. This power of the architectural illusion is so strong that, if we project a perspective onto a vault from below upwardly, not from a projection centre, but following the normal to the impost level, like Tassi does at Palazzo Pallavicini Rospigliosi, we still obtain a strong and happy effect of depth, totally in derogation to the restricted sight, sight that, in this case, simply is not possible.

In conclusion, we believe that a deeper study of the restricted sight, of its limits and its effects, more thorough than this, can eventually free perspective, and its history, from some false opinions that derive from an uncritical reading of Panofsky's famous work. At the same time, this study can unveil which parts of the Panofskyan philosophy that are still entirely valid and which parts that have to be revised in the light of the most recent conquests of science.

ACKNOWLEDGMENTS

The Authors show their gratitude to Professor Mrs. Rita Binaghi for her enlightening translation of several passages of Panofsky's text.

REFERENCES

- [1] Brownson, C. (1981). "Euclid's Optic and its Compatibility with Linear Perspective." *Archive for History of Exact Sciences* (24), 165-194.
- [2] Damisch, H. (1987). *L'origine de la perspective*, Paris: Flammarion.
- [3] Euclide. (2007). *Euclide - Tutte le Opere*, F. Acerbi, F. Acerbi, translator: Bompiani.
- [4] Francesca, P. d. (1899). *Petrus pictor burgensis "de prospectiva pingendi"*, nach dem Codex der königlichen Bibliothete zu Parme nebst deutscher Übersetzung zum Erstenmale veröffentlicht von Dr. C. Winterberg, C. Winterberg, Strassburg: J.H. Ed. Heitz.
- [5] Francesca, P. d. (1984). *De prospectiva pingendi*, G. Nicco Fasola, Firenze: Le Lettere.
- [6] Gioseffi, D. (1957). *Perspectiva artificialis - Per la storia della prospettiva - Spigolature e appunti*.
- [7] Kemp, M. (1990). *The science of art: optical themes in western art from Brunelleschi to Seurat*, New Haven: Yale University Press.
- [8] Kemp, M. (1997). *Behind the picture: art and evidence in the Italian Renaissance*, New Haven: Yale University Press.
- [9] Kemp, M. (1999). "Prospettiva e significato: illusione, allusione e collusione", in *M. Z. Wallace, L., (ed.), Immagine e verità: per una storia dei rapporti tra arte e scienza*. Milano: il Saggiatore.
- [10] Kilpatrick, F. P. (1961). *Explorations in Transactional Psychology*, New York: New York University Press.
- [11] Migliari, R. (2005). "Ha la prospettiva un futuro? (Has Man a future?)", *Ikhmos Analisi grafica e storia della rappresentazione* - 2005. SIRACUSA: Lombardi Editori, pp. 133-160.

- [12] Migliari, R. (2005). "La prospettiva e Panofsky - Panofsky and Perspective." *Disegnare Idee Immagini*, 31, 28-43.
- [13] Migliari, R. (2005). "La prospettiva e Panofsky - Panofsky and Perspective." *Disegnare Idee Immagini*, 31, 28-43.
- [14] Migliari, R. (2006). "La digradazione delle grandezze apparenti nella prospettiva degli antichi", *Le vie dei mercanti - Disegno come topologia della mente*. FIRENZE: Alinea Editrice, pp. 215-224.
- [15] Migliari, R. "Pompei: un trattato antico di prospettiva." Presented at *Le vie dei mercanti - Rappresentare la conoscenza*. Pompei, NAPOLI.
- [16] Monte, G. d. (1984). I sei libri della prospettiva di Guidobaldo dei marchesi del Monte dal latino tradotti interpretati e commentati da Rocco Sinisgalli, R. Sinisgalli, Roma: "L'Erma" di Bretschneider Editrice.
- [17] Panofsky, E. (1991). Perspective as symbolic form, New York Cambridge, Mass.: Zone Books - Distributed by the MIT Press.
- [18] Pozzo, A. (1693). *Perspectiva pictorum et architectorum* Andreae Putei, Roma,: J. J. Komarek.
- [19] Pozzo, A. (1700). *Perspectiva pictorum et architectorum* Andreae Putei ... pars prima[-pars secunda] in quâ docetur modus expeditissimus delineandi optice omnia quæ pertinent ad architecturam, Romæ: Ex typographia Antonii de Rubeis.
- [20] Sartoris, A. (1935). *Gli elementi dell'architettura funzionale*, Milano,: U. Hoepli.
- [21] Sartoris, A. (1949). *Introduzione alla architettura moderna*, Milano,: Hoepli.
- [22] Sartoris, A. (1982). *Progetti e assonometrie di Alberto Sartoris*, M. p. D. A. P. Cortes., Roma: Officina.
- [23] Sartoris, A. (1986). *L'actualité du rationalisme*, Paris: Bibliothèque des arts.
- [24] Somaini, A. (2004). *Rappresentazione prospettica e punto di vista : da Leon Battista Alberti ad Abraham Bosse*, Milano: CUEM.
- [25] Taton, R. (1951). *L'œuvre mathématique de G. Desargues*, Paris: Presses universitaires de France.
- [26] Zevi, B. (1973). *Il linguaggio moderno dell'architettura*. Guida al codice anticlasico, Torino,: G. Einaudi.

ABOUT THE AUTHORS

1. Riccardo Migliari is Full Professor in Fundamentals and Applications of Descriptive Geometry since 1990 at the Faculty of Architecture at the 'La Sapienza' University in Rome. He is assiduously engaged in research, particularly in the areas of Descriptive Geometry and of Representation and Instrumental Survey of Architecture. He directed, as Scientific Manager, the architectural survey of Coliseum in Rome during the preliminary studies for the restoration of the monument undertaken by the Archaeological Superintendence of Rome. From year 2003 he deals, in particular, with the renewal of the studies on the scientific representation of the space, within in the evolutionary picture of the descriptive geometry, from the projective theory to the digital theory and from the graphical applications to the digital modelling.
2. Jessica Romor, architect, Ph.D. at the Department of History, Representation and Restoration of Architecture, Sapienza University of Rome, is interested in traditional and digital representation, in particular for what concern the history of perspective and its applications. She collaborate to the courses of Representation and Descriptive geometry at the Faculty of Architecture at the 'La Sapienza' University in Rome.

POINT IN POLYGON VIA EPI-HYPO GRAPHS, HOMOTOPY AND HOPF'S DEGREE THEOREM

Shriprakash SINHA, Luca NANETTI, Remco J. RENKEN and Gert J. ter HORST
Neuroimaging Center, University Medical Center Groningen, The Netherlands

ABSTRACT: The current work revisits the point in polygon problem by providing a novel solution that explicitly employs the properties of *epigraphs* and *hypographs*. Using concepts of epigraphs and hypographs, this manuscript provides a new definition of *inaccessibility* and *inside*, to accurately specify the meaning of inclusion of a point within or without a polygon. Via Poincaré's ideas on *homotopy* and *Hopf's Degree Theorem* from topology, a relationship between *inaccessibility* and *inside* is established and it is shown that consistent results are obtained for peculiar cases of both non-intersecting and self-intersecting polygons while investigating the point inclusion test w.r.t a polygon. Through illustrative examples, the novel method addresses the issues of • ambiguous solutions given by the cross over for both non-intersecting and self-intersecting polygons and • a point being labeled as multi-ply inside a self-intersecting polygon by the winding number rule, by providing an unambiguous and singular result for both kinds of polygons. The proposed solution bridges the gap between cross-over and winding number rule for complex cases.

Keywords: Point, Polygon, Epigraph, Hypograph, Homotopy, Hopf's Degree Theorem

1. INTRODUCTION

Given a polygon \mathcal{P} or the vertices of the polygon, say $(x_1, y_1), (x_2, y_2) \dots (x_n, y_n)$, it is desired to know whether a sample point $\mathcal{S} (x_0, y_0)$ lies within \mathcal{P} . *The status of a point \mathcal{S} with respect to a polygon \mathcal{P} , termed as inside, needs to be defined properly.* This appears to be crucial in order to retrieve unambiguous results not only for self-intersecting polygons but also for non-intersecting ones in a 2D Cartesian plane.

In this manuscript, a new definition of *inaccessibility* and *inside* has been proposed, aiming to clarify the semantical meaning of *a point being inside (or outside) a polygon*. Later in the manuscript, it is shown that the newly proposed definitions form a bridge between the two well known definitions of Cross Over and the Winding Number Rule. This combined effect of both the existing definitions helps resolve many of the rare issues of solving point in polygon problem for simple as well as self-intersecting polygons.

Cross Over ([18], [4], [5], [6], [7]) states that if a semi infinite line (ray) drawn from \mathcal{S} cuts \mathcal{P} , odd number of times, then the point is inside the polygon. However, there are a few issues regarding this definition. Depending on the **orientation** of the ray from the query point, odd or even number of intersections can be obtained, if the ray passes through vertices. This gives rise to ambiguous results for the same point with different rays at different orientation. In other words, the outcome is a **non-singular** function of the line's direction.

To resolve this issue, a prevalent solution is shifting of the ray infinitesimally ([19], [20]). Here the solution may change drastically depending on the **direction of the shift** and thus the direction of the shift is often fixed. This solution is palliative as the issue is considered to arise rarely and ambiguous results can still be found. A solution to resolve the issue could be to repeat the cross over multiple times until it is found

that the point lies inside the polygon; however, exactly how many repetitions are necessary to prevent a reliable result is unknown leading to **non determinism**.

Winding Number Rule ([7], [3], [1]) states that the number of times one loops around \mathcal{S} while traversing \mathcal{P} before reaching the starting point on the polygon shows whether the point is inside the polygon or not. So a number ℓ greater than one can mean that the point is ℓ times inside the polygon. In this paper, this is considered as an issue because if a point lies inside a polygon once, it lies forever. Thus $\ell > 1$ depicts the idea of **redundancy**. Finally, the Cross Over and the Winding Number Rule algorithms are known to provide different solutions in specific cases of self-intersecting polygons.

The solution provided tries to address these problems by deciding upon the 'inside/outside' status of point with respect to a polygon from a different perspective. [2] proposed to take the decision using a binary coded coordinate system and parity counting of the number of intersections of the polygon with an infinite vector. The algorithm presented here differs from the former in using \mathcal{S} as a reference point. The location of \mathcal{S} does not depend on the coordinate system, but it helps in forming a line (in any orientation) such that it cuts the polygon at different intersection points. For the sake of simplicity, the case where the line is horizontal is presented. The generalization only requires the rotation of the reference (i.e. horizontal-vertical) Cartesian system and the transformation of the coordinate points of the polygon with respect to the rotated system.

The first infinite line is used as a reference to dismember the polygon into sections. These sections (later defined as chains) are then classified as valid or invalid based on definitions of epigraphs or hypographs that may or may not contain \mathcal{S} , respectively. A second line, orthogonal to the first, is used to sort these sections of the polygon that contain the sample point. The sections are then paired and taken out. For each pair the location of the point within the two sec-

tions, is checked. The process of checking continues until the algorithm runs out of pairs to be checked. This iterative procedure of elimination of sections that do not contain the point and final repetitive checking within the pairs of remaining sections, of the polygon, represents the core difference from [2], and is explained in detail in the following section.

A more in-depth theoretical approach to show the correctness of the proposed method follows in section 3, accompanied by a comparison with the Cross Over (section 3.3) and with the Winding Number Rule (section 3.4). The algorithm's computational time scale is discussed in a dedicated section (4) in the present study, followed by results and discussion (section 5) and conclusion (section 6).

2. ALGORITHM IMPLEMENTATION

The algorithm can be described in simple terms as follows. Given the sample point \mathcal{S} with coordinates (x_0, y_0) , a horizontal line $y = y_0$ (i.e. \mathcal{L}_h) is drawn through \mathcal{S} to cut polygon \mathcal{P} at q locations $\{(x_1^{int}, y_0), \dots, (x_q^{int}, y_0)\}$. This breaks the polygon into q chains.

Definition 1. *A chain is a series of connected edges of the polygon whose starting and ending points lie on \mathcal{L}_h , that passes through \mathcal{S} . Mathematically, a chain is a function f_c , with a closed domain defined by the a pair of consecutive intersection points on \mathcal{L}_h here referred by starting and ending points and a range that is the union of the consecutive edges of the polygon comprised between the starting and ending points on \mathcal{L}_h .*

Each chain is then checked for whether its two endpoints contain the test point between them; if not, the chain is discarded. Discarded chains are labeled as *invalid chains* and those kept for further consideration are referred to as *valid chains*. This classification is based on the satisfaction of criterion dependent on concepts of epigraph and hypographs. The definition of these concepts are as follows:

Definition 2. The epigraph of a function (chain) $f_c : \mathcal{R}^n \rightarrow \mathcal{R}$ is a set of points that lie on or above the graph under consideration, such that $\text{epi}(f_c) = \{(x, t) : x \in \mathcal{R}^n, t \in \mathcal{R}, f_c(x) \leq t\}$ is a subset of \mathcal{R}^{n+1} .

Definition 3. The hypograph of a function (chain) $f_c : \mathcal{R}^n \rightarrow \mathcal{R}$ is a set of points that lie on or below the graph under consideration, such that $\text{hypo}(f_c) = \{(x, t) : x \in \mathcal{R}^n, t \in \mathcal{R}, f_c(x) \geq t\}$ is a subset of \mathcal{R}^{n+1} .

Here n equals 2 which is the cardinality of a point in Cartesian plane. The remaining valid chains are then tested for intersection with a vertical line $x = x_0$ through \mathcal{S} . The intersections found are sorted by height, and paired up. If the test point is not between a pair, it is outside. This criterion of containment is checked via the definitions of affine sets and affine combination as follows:

Definition 4. A set $\mathcal{S} \subseteq \mathcal{R}^n$ is an affine set, if for any two points $x_i, x_j \in \mathcal{S}$ ($j > i$) and $\theta \in [0, 1]$, $\theta x_i + (1 - \theta) x_j \in \mathcal{S}$.

Definition 5. An affine combination of $x_i, x_j \in \mathcal{R}$ are a set of points of the the form $\theta_i x_i + \theta_j x_j$, where $\theta_i + \theta_j = 1$.

These definitions and notations, as well as a few others, are adopted from [8]. It should be noted that the vertices of the polygon \mathcal{P} are arranged in order of traversal, starting from any vertex. The traversal order can be in any one direction. Another requisite is that the edges are traversed only once. This is useful in avoiding multiple loops that may occur in cases of self-intersecting polygons.

If \mathcal{S} lies out of the bounding box of the polygon, it is considered outside \mathcal{P} and no further processing is done. Lastly, if the sample point is one of the vertices of the polygon, then it is considered to be in the polygon. This final point is assumed as the proposed algorithm would reach the same conclusion at the expense of computational time. In the theoretical proof it will be

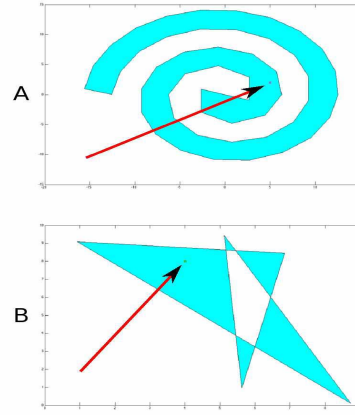


Figure 1: Query point location in (A) non-intersecting and (B) self-intersecting polygon.

shown that this case is true and from implementation point of view the idea holds correct.

What follows is a step-by-step explanation of how the algorithm works; pictorial representations will help in clarifying each step. The examples will regard closed polygons, both self-intersecting, and non-intersecting. Figures 1 and 2 show the polygons with the sample point being tested at different locations.

2.1 Intersecting the \mathcal{P}

It is known that \mathcal{P} is an ordered series of vertices $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$, starting from (x_1, y_1) such that the ending point after the traversal has coordinates $(x_{n+1}=1, y_{n+1}=1)$. Given \mathcal{P} , the first step is to draw \mathcal{L}_h through the sample point \mathcal{S} (x_0, y_0) , such that it intersects the polygon at certain points. As observed before, the line \mathcal{L}_h being horizontal does not imply any loss of generality in the proposed solution for the point in polygon problem.

The intersection point is obtained by computing the coordinate values of the common point between \mathcal{L}_h and a straight edge extending from (x_i, y_i) to (x_{i+1}, y_{i+1}) (henceforth

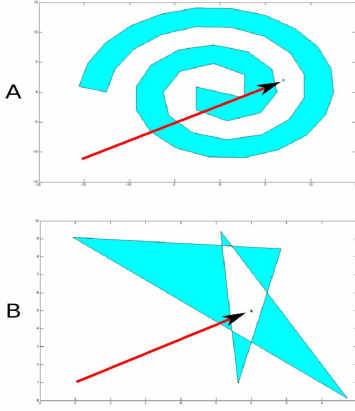


Figure 2: Query point location in (A) non-intersecting and (B) self-intersecting polygon.

$(x_i, y_i), (x_{i+1}, y_{i+1})$). The slope and the constant of the former is 0 and y_0 , and that of latter is $m_i = \frac{y_{i+1} - y_i}{x_{i+1} - x_i}$ and $c_i = y_i - m_i x_i$. Here i and $i + 1$ are indices to any pair of consecutive vertices on \mathcal{P} . Let the coordinate values of this common point be (x^{int}, y_0) . Solving the algebraic equation between the two straight lines yield:

$$x^{int} = \frac{y_0 - c_i}{m_i} \quad (1)$$

In this case when the infinite line \mathcal{L}_h is horizontal, the y coordinate of the intersection point equals to the y coordinate of \mathcal{L}_h , i.e. y_0 . Once the intersecting point with coordinate (x^{int}, y_0) is obtained, the algorithm checks it's inclusion on $\overline{(x_i, y_i), (x_{i+1}, y_{i+1})}$ of \mathcal{P} , for further processing. If the criterion of inclusion is satisfied, the intersection point is stored as a new vertex point that needs to be appended to the pre-existing list of vertices of the polygon at a later stage. This inclusion criterion is put into affect by the use of the affine combination property in definition 5.

Three different cases arise depending on the slope of $\overline{(x_i, y_i), (x_{i+1}, y_{i+1})}$:

- $m_i = \pm\infty$: If the edge is a vertical line, the

point (x^{int}, y_0) lies on the line. This is because $x^{int} = x_i = x_{i+1}$.

- $m_i = 0$: If edge is a horizontal line itself, then the intersection point (x^{int}, y_0) is considered to lie outside the range of (x_i, y_0) and (x_{i+1}, y_0) . This is because, if $(x^{int}, y_0) \in [(x_i, y_0), \text{ and } (x_{i+1}, y_0)]$, then there are infinitely many values that could be assigned to x^{int} . To avoid random selection of any point within the above range, the point (x^{int}, y_0) is considered to lie outside the range. Also, if (x_0, y_0) lie on a horizontal edge, it is still considered outside the range for further processing.
- $m_i \in \mathcal{R} - \{0, \pm\infty\}$: Finally, this being the simplest case, it is easy to compute whether (x^{int}, y_0) lies on the line between the given points using definition 5.

It is important to note that this process of finding intersection points and their subsequent inclusion into the pre-existing vertex list of \mathcal{P} , based on satisfaction of affine combination property, is iterative in nature: that is, for each pair of (x_i, y_i) and (x_{i+1}, y_{i+1}) as i iterates through values 1 to n , the intersection points are computed between $\overline{(x_i, y_i), (x_{i+1}, y_{i+1})}$ and \mathcal{L}_h . Next, for each edge $\overline{(x_i, y_i), (x_{i+1}, y_{i+1})}$ and its corresponding intersection point with \mathcal{L}_h , it is verified whether the intersection point lies in between (x_i, y_i) and (x_{i+1}, y_{i+1}) . If yes, then the points are stored. The stored points are then appended to the list of pre-existing vertices \mathcal{P} , such that the traversal order remains unaffected. In other words, if $(x^{int}, y_0) \in \overline{(x_i, y_i), (x_{i+1}, y_{i+1})}$ then a particular subsequence of the traversal order becomes: $(x_i, y_i) \rightarrow (x^{int}, y_0) \rightarrow (x_{i+1}, y_{i+1})$.

Figures 3 and 4 show \mathcal{L}_h passing through \mathcal{P} and intersecting the polygon at different edges. The location of \mathcal{P} is indicated by the red arrow while the new vertices at the intersection of \mathcal{L}_h with edges of \mathcal{P} are pointed by the blue arrows. Some of the newly added vertices may lie very close to the pre-existing vertices e.g. in

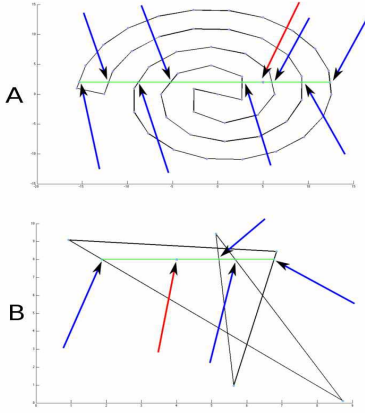


Figure 3: Finding intersection points on the polygon. Location of \mathcal{S} indicated via red arrow in (A) non-intersecting and (B) self-intersecting polygon. Blue arrows point to newly found intersection points. The green line depicts \mathcal{L}_h i.e $y = y_0$.

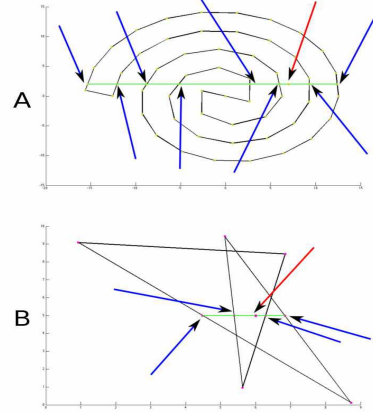


Figure 4: Finding intersection points on the polygon. Location of \mathcal{S} indicated via red arrow in (A) non-intersecting and (B) self-intersecting polygon. Blue arrows point to newly found intersection points. The green line depicts \mathcal{L}_h i.e $y = y_0$.

the range of $\pm 10^{-5}$ (i arbitrary units) or smaller. The algorithm removes these old vertices from the list that lie in such close range, but stores the newly added vertices with their coordinates (x^{int}, y_0) separately. Two reasons arise for executing this step:

- To avoid further computations that may involve floating point precision of the order smaller than or equal to $\pm 10^{-5}$.
- Retention of newly appended vertices with coordinates (x^{int}, y_0) will be later used for searching chains whose epi/hypo-graph may contain \mathcal{S} .

In this paper, the tolerance range is $(\pm 10^{-5})$ is an arbitrarily assigned value.

2.2 Decomposition of Polygon into Valid and Invalid Chains

The new vertices with coordinates (x_j^{int}, y_0) (where $j \in \{1, \dots, m\}$) and the sample point (x_0, y_0) form the basis for the next steps. From

definition 2 it is known that a point belongs to the epigraph (hypograph) if it lies on or above (below) the function under consideration. To use the mentioned properties, the polygon \mathcal{P} is decomposed into chains. These chains would then be tested for convexity or concavity with respect to \mathcal{L}_h in the following way: One of the newly added vertex on \mathcal{P} (with the y coordinate being y_0) is picked up as the starting vertex. A traversal order is chosen randomly and is followed until the starting point is reached again.

As the traversal is done from one intersecting vertex to the other, the polygon gets decomposed into subsets of consecutive edges, thus forming chains. Each chain may contain more than one original vertex of the polygon excluding the starting and ending vertex of the chain. These chains lie either above or below \mathcal{L}_h i.e $y = y_0$. The chains are further classified as *valid* or *invalid* using the definitions in 2 and 3. In non mathematical terms, if the starting and ending vertices of a chain are on different side and not on the same

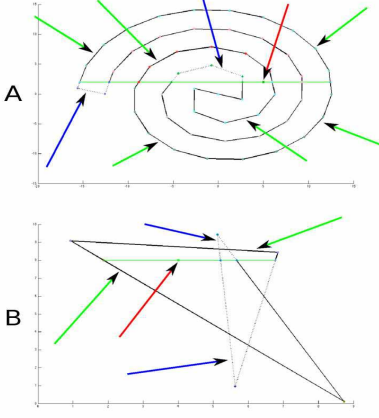


Figure 5: (A) Non-intersecting and (B) self-intersecting polygon decomposed into chains. Green and blue arrows indicate the valid and invalid chains, respectively. Location of query point w.r.t the polygon is shown via the red arrow.

side with respect to \mathcal{S} on \mathcal{L}_h , then the chain is a valid one: else the chain is invalid. The invalid chains are discarded and the valid ones are stored with their starting and ending vertices along with coordinates of vertices on the chain.

Figures 5 and 6 show the valid and invalid chains pointed by the green and blue arrows respectively. The solid lines represent the valid chains and the dotted lines represent the invalid chains. The red arrow indicates the sample point's location in each of the figures.

2.3 Chain Intersection

Hitherto, it is known that the x coordinate of \mathcal{S} lies in the epi/hypo-graph of the valid chains. To decide if the point lies inside or is inaccessible with respect to a polygon under consideration, what needs to be tested is whether the y coordinate of \mathcal{S} lies within any two nearest valid chains. The rationale behind doing these steps will be elucidated a little later, but before that it is important to define what the nearest valid chains

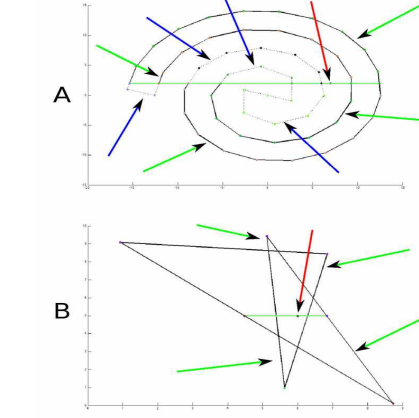


Figure 6: (A) Non-intersecting and (B) self-intersecting polygon decomposed into chains. Green and blue arrows indicate the valid and invalid chains, respectively. Location of query point w.r.t the polygon is shown via the red arrow.

mean:

Definition 6. Chains \mathcal{C}_i and \mathcal{C}_j are nearest valid chains if:

- the $\text{epi}(f_{\mathcal{C}_i}) \subset \text{epi}(f_{\mathcal{C}_u}) \forall u \in 1, \dots, q$ chains below \mathcal{S} such that $(x_0, y_0) \in \text{epi}(f_{\mathcal{C}_u})$.
- the $\text{hypo}(f_{\mathcal{C}_j}) \subset \text{hypo}(f_{\mathcal{C}_v}) \forall v \in 1, \dots, q$ chains above \mathcal{S} such that $(x_0, y_0) \in \text{hypo}(f_{\mathcal{C}_v})$.

After all the valid chains have been retained, a similar procedure (as before) of intersecting the valid chains using the vertical line $x = x_0$ (\mathcal{L}_v) that passes through \mathcal{S} , is executed. For each valid chain, defined by a set of vertices (x_i, y_i) (such that $i \subseteq \{1, \dots, n\}$), points of intersection are computed between the edges of the chain and \mathcal{L}_v . The process of evaluating the intersection points follows: The slope of the straight edge joining two vertices on the chain is $m_i = \frac{y_{i+1} - y_i}{x_{i+1} - x_i}$ and the constant is $c_i = y_i - m_i x_i$. Here, i and $i+1$

are consecutive vertices on the valid chain. The value of y coordinate of the point of intersection is computed as follows:

$$y^{int} = m_i x_0 + c_i \quad (2)$$

Once the coordinates of the intersection point (x_0, y^{int}) is obtained for an edge, a test is conducted to find whether the intersection point lies in between the edge $(x_i, y_i), (x_{i+1}, y_{i+1})$. This is achieved using the affine combination property in definition 5. Three cases may arise, that need to be considered:

- $m_i = \pm\infty$: If edge is a vertical line, the intersecting point (x_0, y^{int}) is considered to lie inside the range of (x_0, y_i) and (x_0, y_{i+1}) . This is because, if the chain crosses $x = x_0$ many times before going from left of \mathcal{S} to the right of \mathcal{S} or vice versa, then there can be infinitely many points on one chain that may be considered as intersection points. That is, it may require the algorithm to store many intersection points for just one chain. To avoid the presence of multiple intersection points on a single chain, the algorithm stores the vertex of one of the vertical edges in a chain.
- $m_i = 0$: If the edge is a horizontal line, the intersecting point (x_0, y^{int}) lies on the line. This is because $x^{int} = x_0$ and $y^{int} = y_i = y_{i+1}$.
- $m_i \in \mathcal{R} - \{0, \pm\infty\}$: Finally, this being the simplest case, it is easy to compute whether (x_0, y^{int}) lies on the line between the given points using definition 5.

This process of finding the intersection point is repeated for all edges (i.e. $\forall i \subseteq \{1, \dots, n\}$) that constitute the valid chain under consideration. Next, for each edge $(x_i, y_i), (x_{i+1}, y_{i+1})$ and its corresponding intersection point with \mathcal{L}_v , it is verified whether the intersection point lies within (x_i, y_i) and (x_{i+1}, y_{i+1}) . It is expected that, after preprocessing of all edges on a chain, there exists only one intersection point between a chain and

\mathcal{L}_v . The reason behind this is to find one common point of intersection in between the valid chain and \mathcal{L}_h . This is done as a necessary step to sort the valid chains on the basis of y coordinate value of the newly found intersection points. The sorted valid chains will further be processed to test the inclusion of \mathcal{S} within \mathcal{P} .

The pictorial representation of the line $x = x_0$ intersecting the valid chains are shown in figures 7 and 8. The green arrows indicate the points of intersection on the valid chains. The valid chains are pointed by the blue arrows. The sample point is indicated via the red arrow. As mentioned earlier, the invalid dotted chains have been removed by the algorithm, in the final stages of the processing.

2.4 Point Inclusion Test

Each of the valid chains now has a point of intersection with \mathcal{L}_v . The algorithm sorts the series of intersection points based on their y coordinates (i.e. y^{int}) that lie on \mathcal{L}_v , thus sorting the valid chains.

It is also known that \mathcal{S} lies either in the epigraph or in the hypograph of any valid chain. The final step to decide whether \mathcal{S} is inside \mathcal{P} is to assess whether the y coordinate of \mathcal{S} is compartmentalized between two nearest valid chains \mathcal{C}_i and \mathcal{C}_{i+1} (if \mathcal{C}_i is below \mathcal{C}_{i+1} after sorting), such that $\mathcal{S} \in \text{epi}(f_{\mathcal{C}_i})$ and $\mathcal{S} \in \text{hypo}(f_{\mathcal{C}_{i+1}})$. This is done as follows:

Considering a pair of y^{int} 's at a time, i.e a pair of chains at a time, it is tested whether $\mathcal{S} = (x_0, y_0)$ is an affine combination of (x_0, y_j^{int}) and (x_0, y_i^{int}) (for $i < j$). If such a pair is found, then the point lies inside the polygon, else it is outside. The only constraint is that the pair of chains or the pair of intersecting y coordinates are mutually exclusive. Thus, if there are $m/2$ pairs, with m being an even number and $y_1^{int}, y_2^{int}, \dots, y_m^{int}$ are the intersecting y values in order, then $(y_1^{int}, y_2^{int}), (y_3^{int}, y_4^{int}), \dots, (y_{m-1}^{int}, y_m^{int})$ are mutually exclusive in the sense that element of one pair cannot be included in any other pair.

The intuitive idea behind this rule is that, in a

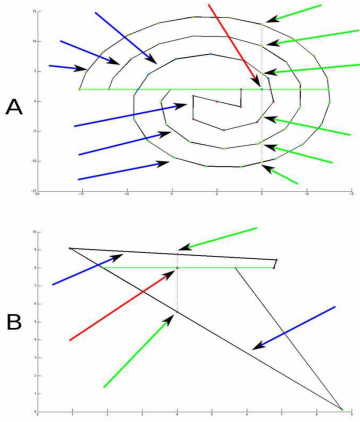


Figure 7: Inclusion test, \mathcal{S} inside \mathcal{P} . Blue arrows: valid chains of (A) non-intersecting and (B) self-intersecting polygon. Green arrows: intersection points with the vertical line \mathcal{L}_v . The point \mathcal{S} (red arrow) is inside the polygon.

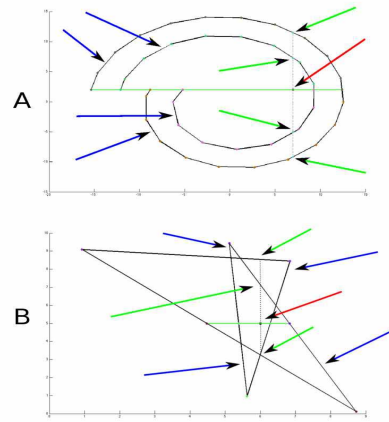


Figure 8: Inclusion test, \mathcal{S} outside \mathcal{P} . Blue arrows: valid chains of (A) non-intersecting and (B) self-intersecting polygon. Green arrows: intersection points with the vertical line \mathcal{L}_v . The point \mathcal{S} (red arrow) is outside the polygon.

pair, if the path of traversal in a chain is moving (say) from left side of \mathcal{S} to the right side, then the traversal in the other chain must move from right side \mathcal{S} to the left (and vice versa). Thus, any pair shall not contain chains from any other pair. These ideas can be seen in figures 7 and 8. The green arrows mark the intersecting points on the valid chains. Testing the sample point \mathcal{S} as an affine combination of two intersection points on \mathcal{L}_v indicates whether the point lies within or without the polygon.

3. INACCESSIBILITY-INSIDE THEOREM

The meaning of inside is viewed from different perspectives via the definitions of the cross over and the winding number rule. This gives rise to contradictory results in peculiar cases of non-intersecting and self-intersecting polygons.

This manuscript proposes new definitions of *inside* and *inaccessibility* of a point \mathcal{S} with respect to a polygon \mathcal{P} . Also, a relation between inaccessibility and inside is proved. It is shown that

consistent results can be obtained if the meaning of the inaccessibility and inside of a polygon is framed correctly, in an abstract sense. It must be noted that the points that lie on vertices are special cases and the definitions of *inside* and *inaccessibility* will have to be slightly modified to take them into account without changing the general, abstract meaning of inside and inaccessibility. Ultimately, two cases that are examined are: a point lying on (1) a vertex and (2) an edge or anywhere else.

3.1 Point on Vertex of Polygon

Definition 7. The inaccessibility $Inacc_{\mathcal{P}}(\mathcal{S})$ of a point \mathcal{S} related to a polygon \mathcal{P} , is the number of valid chains that need to be broken by a line passing through \mathcal{S} such that the line reaches outside the bounding box of the polygon.

$$Inacc_{\mathcal{P}}(\mathcal{S}) = \begin{cases} \mathcal{N}, & \mathcal{N} \neq 0 \text{ valid chains} \\ & \text{to be broken} \\ 0, & \text{otherwise} \end{cases}$$

Definition 8. The status of a point \mathcal{S} related to a polygon \mathcal{P} , that is $Inside_{\mathcal{P}}(\mathcal{S})$, is the existence of a chain \mathcal{C} such that $\mathcal{S} \in epi(f_{\mathcal{C}})$ or $\mathcal{S} \in hypo(f_{\mathcal{C}})$.

$$Inside_{\mathcal{P}}(\mathcal{S}) = \begin{cases} 1, & \text{if } \mathcal{S} \in epi(f_{\mathcal{C}}) \\ & \text{or } \mathcal{S} \in hypo(f_{\mathcal{C}}) \\ 0, & \text{otherwise} \end{cases}$$

These two definitions form the basis of the two theorems below.

Theorem 1. Given that a point \mathcal{S} is inside as well as inaccessible polygon w.r.t \mathcal{P} , then the following statements: (1) $Inside_{\mathcal{P}}(\mathcal{S}) = 1$, (2) $Inacc_{\mathcal{P}}(\mathcal{S}) = \mathcal{N}$, are logically equivalent, i.e. $Inside_{\mathcal{P}}(\mathcal{S}) = 1$ iff $Inacc_{\mathcal{P}}(\mathcal{S}) = \mathcal{N}$

Proof. (a) If $Inacc_{\mathcal{P}}(\mathcal{S}) = \mathcal{N}$ then $Inside_{\mathcal{P}}(\mathcal{S}) = 1$

Given that $Inacc_{\mathcal{P}}(\mathcal{S}) = \mathcal{N}$, there exists \mathcal{N} valid chains that need to be broken according to definition 7. It is known that a chain is valid when either it's epigraph or hypograph contains \mathcal{S} . This existence of \mathcal{N} valid chains imply that $\mathcal{S} \in \{epi(f_{\mathcal{C}_k}), hypo(f_{\mathcal{C}_k})\} \forall k \in \{1, \dots, \mathcal{N}\}$. But this is the definition of status of \mathcal{S} related to \mathcal{P} , i.e. $Inside_{\mathcal{P}}(\mathcal{S}) = 1$ or $Inside_{\mathcal{P}}(\mathcal{S}) \in \{1\}$.

(b) If $Inside_{\mathcal{P}}(\mathcal{S}) = 1$ then $Inacc_{\mathcal{P}}(\mathcal{S}) = \mathcal{N}$

Given $Inside_{\mathcal{P}}(\mathcal{S}) = 1$ implies that $\mathcal{S} \in \{epi(f_{\mathcal{C}}), hypo(f_{\mathcal{C}})\}$ for a chain \mathcal{C} in f . Thus chain \mathcal{C} is a valid chain, as it contains the point \mathcal{S} . In order for \mathcal{S} to be inaccessible, there must exist atleast 1 valid chain in \mathcal{P} that needs to be broken. Since \mathcal{C} is one such chain and the only chain that contains \mathcal{S} , the inaccessibility order of \mathcal{S} related to \mathcal{P} in $Inacc_{\mathcal{P}}(\mathcal{S}) = 1$ or $Inacc_{\mathcal{P}}(\mathcal{S}) \in \{1\}$.

If \mathcal{S} is a vertex such that it is an intersection point of two or more lines of a polygon, then all chains that have their epigraph or hypograph

containing \mathcal{S} , are valid, since it requires \mathcal{N} (if \mathcal{N} is the number of valid chains) chains to be broken. \square

Pictures will help the reader to get acquainted with the practical consequences of the previous theorem. Figure 9 shows three different polygons with \mathcal{S} as the point under consideration. The polygon in figure 9.(A) has four chains that contain \mathcal{S} namely (a) \mathcal{STU} (b) $UV\mathcal{S}$ (c) $\mathcal{S}WX$ and (d) $XY\mathcal{S}$, which are valid. Thus by theorem 1, $Inside_{\mathcal{P}}(\mathcal{S}) = 1$ and $Inacc_{\mathcal{P}}(\mathcal{S}) = 4$. Thus \mathcal{S} lies inside the polygon. Similarly, for figure 9.(B) there is one chain $\mathcal{STU}\mathcal{S}$ which is valid as it contains the point \mathcal{S} . Thus $Inside_{\mathcal{P}}(\mathcal{S}) = 1$ and $Inacc_{\mathcal{P}}(\mathcal{S}) = 4$. For the case of figure 9.(C) there exists two chains that contain \mathcal{S} i.e. (a) $\mathcal{STU}\mathcal{S}$ and $\mathcal{S}VW\mathcal{S}$ which are valid. So $Inside_{\mathcal{P}}(\mathcal{S}) = 1$ and $Inacc_{\mathcal{P}}(\mathcal{S}) = 2$.

Since it is always true that \mathcal{S} is inside \mathcal{P} when it lies on one of its vertices, it is a reasonable strategy to check this first, avoiding further computations. Thus in the current implementation of the algorithm, the sample point is always first checked against the vertices of the polygon in order to know if it belongs to \mathcal{P} .

Theorem 2. Given a point \mathcal{S} is not inside as well as not inaccessible w.r.t polygon \mathcal{P} , then the following statements: (1) $Inside_{\mathcal{P}}(\mathcal{S}) = 0$ and (2) $Inacc_{\mathcal{P}}(\mathcal{S}) = 0$, are logically equivalent, i.e: $Inside_{\mathcal{P}}(\mathcal{S}) = 0$ iff $Inacc_{\mathcal{P}}(\mathcal{S}) = 0$

Proof. (a) If $Inacc_{\mathcal{P}}(\mathcal{S}) = 0$ then $Inside_{\mathcal{P}}(\mathcal{S}) = 0$

Given that $Inacc_{\mathcal{P}}(\mathcal{S}) = 0$, there exists no valid chains that need to be broken according to definition 7. This means that $\mathcal{S} \notin \{epi(f_{\mathcal{C}_k}), hypo(f_{\mathcal{C}_k})\} \forall k$ chains in \mathcal{P} . Since no chain exists whose epigraph or hypograph contains \mathcal{S} , the status of \mathcal{S} related to \mathcal{P} is $Inside_{\mathcal{P}}(\mathcal{S}) = 0$ or $Inside_{\mathcal{P}}(\mathcal{S}) \in \{0\}$.

(b) If $Inside_{\mathcal{P}}(\mathcal{S}) = 0$ then $Inacc_{\mathcal{P}}(\mathcal{S}) = 0$

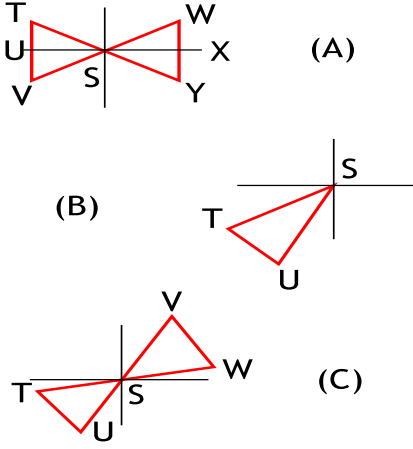


Figure 9: Polygons with locations of the point \mathcal{S} .

Given $Inside_{\mathcal{P}}(\mathcal{S}) = 0$ implies that $\mathcal{S} \notin \{epi(f_{\mathcal{C}_k}), hypo(f_{\mathcal{C}_k})\} \forall k$ chains in \mathcal{P} . This means no valid chains exist in \mathcal{P} that need to be broken. Thus the inaccessibility of \mathcal{S} related to \mathcal{P} is zero, i.e. $Inacc_{\mathcal{P}}(\mathcal{S}) \in \{0\}$, which is the desired result. \square

Cases for theorem 2 are simple and depicted in figure 10. Figure 10 shows two different polygons with \mathcal{S} as the point under consideration. The polygon in figure 10.(A) has four chains that do not contain \mathcal{S} namely (a) RTU (b) UVR (c) RWX and (d) XYR, which are invalid. Thus by theorem 2, $Inside_{\mathcal{P}}(\mathcal{S}) = 0$ and $Inacc_{\mathcal{P}}(\mathcal{S}) = 0$. Thus \mathcal{S} lies outside the polygon. Similarly, for figure 10.(B) there exists two chains that do not contain \mathcal{S} i.e. (a) RTUR and RVWR, which are invalid. So $Inside_{\mathcal{P}}(\mathcal{S}) = 0$ and $Inacc_{\mathcal{P}}(\mathcal{S}) = 0$.

3.2 Point not on Vertex of Polygon

Definition 9. The inaccessibility $Inacc_{\mathcal{P}}(\mathcal{S})$ of a point \mathcal{S} related to a polygon \mathcal{P} , is the number of pairs of valid chains that need to be broken and/or ignored by a line passing through \mathcal{S} such that the line reaches outside the bounding box of

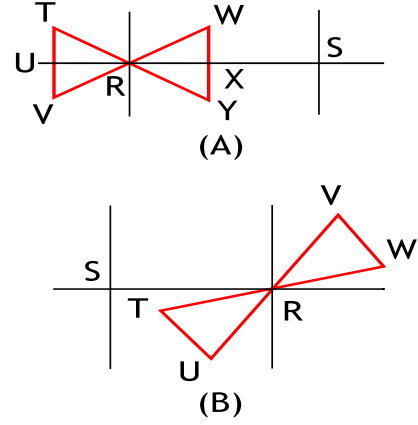


Figure 10: Polygons with locations of the point \mathcal{S} .

the polygon.

$$Inacc_{\mathcal{P}}(\mathcal{S}) = \begin{cases} 1, & \text{a pair of chains need to be broken} \\ \mathcal{N}, & \mathcal{N} \neq 0 \text{ pairs of chains to be ignored} \\ 1 + \mathcal{N}, & \text{a pair to be broken and } \mathcal{N} \text{ pairs to be ignored} \end{cases}$$

Definition 10. The status of a point \mathcal{S} related to a polygon \mathcal{P} , that is $Inside_{\mathcal{P}}(\mathcal{S})$, is the existence of a pair of chains \mathcal{C}_i and \mathcal{C}_j such that $\mathcal{S} \in epi(f_{\mathcal{C}_i})$ and $\mathcal{S} \in hypo(f_{\mathcal{C}_j})$ ($i < j$).

$$Inside_{\mathcal{P}}(\mathcal{S}) = \begin{cases} 1, & \text{pairs of chains } \mathcal{C}_i \text{ and } \mathcal{C}_j, \\ & \text{s.t. } \mathcal{S} \in epi(f_{\mathcal{C}_i}) \text{ and } \mathcal{S} \in hypo(f_{\mathcal{C}_j}) \\ 0, & \text{otherwise} \end{cases}$$

The theorems proving the relation between inaccessibility and inside of a polygon is as follows:

Theorem 3. *Given a point \mathcal{S} is inside as well as inaccessible w.r.t polygon \mathcal{P} , then the following statements (1) $Inside_{\mathcal{P}}(\mathcal{S}) = 1$ and (2) $Inacc_{\mathcal{P}}(\mathcal{S}) = 1$ or $Inacc_{\mathcal{P}}(\mathcal{S}) = 1 + \mathcal{N}$, are logically equivalent, i.e.: $Inside_{\mathcal{P}}(\mathcal{S}) = 1$ iff $Inacc_{\mathcal{P}}(\mathcal{S}) = 1$ or $Inacc_{\mathcal{P}}(\mathcal{S}) = 1 + \mathcal{N}$*

Proof. (a) If $Inacc_{\mathcal{P}}(\mathcal{S}) = 1$ or $Inacc_{\mathcal{P}}(\mathcal{S}) = 1 + \mathcal{N}$ then $Inside_{\mathcal{P}}(\mathcal{S}) = 1$.

Given $Inacc_{\mathcal{P}}(\mathcal{S}) \in \{1, 1 + \mathcal{N}\}$ implies that there exist a pair of valid chains in \mathcal{P} that need to be broken and/or \mathcal{N} pairs of valid chains that need to be ignored. A valid chain by definition is one whose epigraph or hypograph contains \mathcal{S} . Taking the general case of $1 + \mathcal{N}$ (if $\mathcal{N} = 0$, $1 + \mathcal{N}$ collapses to 1), there are $2 \times (1 + \mathcal{N})$ valid chains such that half lie above/on \mathcal{S} and the rest half lie below/on \mathcal{S} . If a vertical line \mathcal{L}_v is drawn such that it cuts the valid chains and \mathcal{S} , then the chains can be sorted according to the value of intersection points in $x = x_0$. Let $\mathcal{C}_1, \dots, \mathcal{C}_{2 \times (1 + \mathcal{N}) - 1}, \mathcal{C}_{2 \times (1 + \mathcal{N})}$ be the sorted order of chains from bottom to top. Taking mutually exclusive consecutive pairs of these valid chains i.e. $(\mathcal{C}_1, \mathcal{C}_2), (\mathcal{C}_3, \mathcal{C}_4), \dots, (\mathcal{C}_i, \mathcal{C}_{i+1}), \dots, (\mathcal{C}_{2 \times (1 + \mathcal{N}) - 1}, \mathcal{C}_{2 \times (1 + \mathcal{N})})$, it is easy to know whether \mathcal{S} is an affine combination of $(x_0, y_{\mathcal{C}_k}^{int})$ and $(x_0, y_{\mathcal{C}_{k+1}}^{int})$, $\forall k \in \{1, 3, 5, \dots, 2 \times (1 + \mathcal{N}) - 1\}$. Here $y_{\mathcal{C}_k}^{int}$ and $y_{\mathcal{C}_{k+1}}^{int}$ are the y coordinates of the intersection points on \mathcal{L}_v and the chains k and $k + 1$, respectively.

Since it is known that atleast one pair of chains need to be broken, a pair of points $(x_0, y_{\mathcal{C}_i}^{int})$ and $(x_0, y_{\mathcal{C}_{i+1}}^{int})$ exists for which $\mathcal{S} (x_0, y_0) \equiv (x_0, \theta y_{\mathcal{C}_i}^{int} + (1 - \theta) y_{\mathcal{C}_{i+1}}^{int})$ for $0 \leq \theta \leq 1$. This implies that \mathcal{C}_i is the nearest chain below/on \mathcal{S} and \mathcal{C}_{i+1} is the nearest chain above/on \mathcal{S} , otherwise \mathcal{S} won't be an affine combination of $(x_0, y_{\mathcal{C}_i}^{int})$ and $(x_0, y_{\mathcal{C}_{i+1}}^{int})$. For the rest of the \mathcal{N} pairs, since \mathcal{S} is not an affine combination of $(x_0, y_{\mathcal{C}_k}^{int})$ and $(x_0, y_{\mathcal{C}_{k+1}}^{int}) \forall k \in \{1, 3, \dots, 2 \times (1 + \mathcal{N}) - 1\} - \{i\}$, these \mathcal{N} pairs of chains can be ignored for further processing or consideration. Since these

nearest chains \mathcal{C}_i and \mathcal{C}_{i+1} are valid also, their epigraph and hypograph contain \mathcal{S} , respectively. This existence of a pair of a valid chains which has to be broken such that $\mathcal{S} \in epi(f_{\mathcal{C}_i})$ and $\mathcal{S} \in hypo(f_{\mathcal{C}_{i+1}})$ implies $Inside_{\mathcal{P}}(\mathcal{S}) = 1$, the status of \mathcal{S} related to \mathcal{P} .

(b) If $Inside_{\mathcal{P}}(\mathcal{S}) = 1$ then $Inacc_{\mathcal{P}}(\mathcal{S}) = 1$ or $Inacc_{\mathcal{P}}(\mathcal{S}) = 1 + \mathcal{N}$

Let \mathcal{P} be a polygon such that $Inside_{\mathcal{P}}(\mathcal{S}) = 1$ implies there exists a pair of chains \mathcal{C}_i and \mathcal{C}_j such that $\mathcal{S} \in epi(f_{\mathcal{C}_i})$ and $\mathcal{S} \in hypo(f_{\mathcal{C}_j})$. Given only these two chains, it is evident that both of them are nearest chains to \mathcal{S} . Let the starting and ending points of \mathcal{C}_i and \mathcal{C}_j be $\{(x_{\mathcal{C}_{is}}^{int}, y_0), (x_{\mathcal{C}_{ie}}^{int}, y_0)\}$ and $\{(x_{\mathcal{C}_{js}}^{int}, y_0), (x_{\mathcal{C}_{je}}^{int}, y_0)\}$, respectively. If a vertical line $x = x_0$ is drawn through (x_0, y_0) it would intersect the chains \mathcal{C}_i and \mathcal{C}_j at $(x_0, y_{\mathcal{C}_i}^{int})$ and $(x_0, y_{\mathcal{C}_j}^{int})$, respectively. Since $(x_0, y_{\mathcal{C}_i}^{int})$ lies below (x_0, y_0) and $(x_0, y_{\mathcal{C}_j}^{int})$ lies above (x_0, y_0) , $\mathcal{S} (= (x_0, y_0))$ is an affine combination of $(x_0, y_{\mathcal{C}_i}^{int})$ and $(x_0, y_{\mathcal{C}_j}^{int})$. Thus \mathcal{S} lies between \mathcal{C}_i and \mathcal{C}_j . Now, if an end point of \mathcal{C}_i is joined with an end point of \mathcal{C}_j and another end point of the former joined to the remaining end point of the latter, then a closed loop is formed such that traversing once from any one point, leads to the same point in the end. Let this loop be \mathcal{P}' . **To check for the inaccessibility, it needs to be known whether \mathcal{P}' can be transformed into \mathcal{P} while retaining all the geometric properties.**

Poincaré's ideas on homotopy ([9], [10], [11] and [12]) state that *geometric objects or simple cases of paths as continuous functions are homotopic if one function can continuously deform into another*. From [14], if \mathcal{X} and \mathcal{Y} be two (topological) spaces and $f, g: \mathcal{X} \rightarrow \mathcal{Y}$ two (continuous) maps of \mathcal{X} and \mathcal{Y} , then formally:

Definition 11. *The maps f and g are homotopic, if there exists a map $\mathcal{F}: \mathcal{X} \times [0, 1] \rightarrow \mathcal{Y}$ such that $\mathcal{F}(x, 0) = f(x)$ and $\mathcal{F}(x, 1) = g(x)$. Such a map \mathcal{F} is called a homotopy between f and g .*

Simplifying the above scenario for 2 dimensional Cartesian planes, f and g are functions in 2D space and \mathcal{F} is a mapping that helps in transformation of f to g . In case such a mapping exists, then \mathcal{F} is a homotopy between the two functions.

To retain geometric properties while transformation, **Hopf's Degree Theorem** [13] states that a loop \mathcal{A} may be continuously deformed into another loop \mathcal{B} , without ever crossing the point \mathcal{D} , if and only if \mathcal{A} and \mathcal{B} have the same winding number round \mathcal{D} . Formally, from [14], the Hopf's degree theorem is stated as follows:

Theorem 4. *If \mathcal{X} is a compact, connected, oriented, n -dimensional manifold without boundary and \mathcal{S}^n the n -sphere, then the two maps $f, g : \mathcal{X} \rightarrow \mathcal{S}^n$ are continuously homotopic if and only if $\deg(f) = \deg(g)$.*

Here the *deg* or *degree* of a map is a *homotopy invariant*, i.e the maps respect the relation of *homotopy equivalence*. Definitions and proofs of concept for *degree* of a map and it being a *homotopy invariant* is beyond the scope of this manuscript. Interested readers are advised to consult [14]. In the current situation f and g are the loops \mathcal{P}' and \mathcal{P} , respectively. As long as the winding number of \mathcal{P}' around \mathcal{S} is same as that of \mathcal{P} around \mathcal{S} , \mathcal{P}' and \mathcal{P} are homotopic i.e \mathcal{P}' can be deformed into \mathcal{P} , by Theorem 4. Without loss of generality, let \mathcal{P}' can be deformed into \mathcal{P} .

Since \mathcal{P} is formed from the two necessary chains \mathcal{C}_i and \mathcal{C}_j which are valid, a pair exists in polygon \mathcal{P} that needs to be *broken*. Thus the minimum inaccessibility of \mathcal{S} related to \mathcal{P} is $Inacc_{\mathcal{P}}(\mathcal{S}) = 1$. If there exists extra pairs of valid chains, then they would be *ignored* from consideration, while checking for the affine combination criteria of \mathcal{S} with respect to the sorted pairs of chains on $x = x_0$. If \mathcal{N} is the minimum number of pairs of valid chains that are *ignored*, then the inaccessibility of \mathcal{S} related to \mathcal{P} is $Inacc_{\mathcal{P}}(\mathcal{S}) = 1 + \mathcal{N}$. Thus $Inacc_{\mathcal{P}}(\mathcal{S}) = \{1, 1 + \mathcal{N}\}$. \square

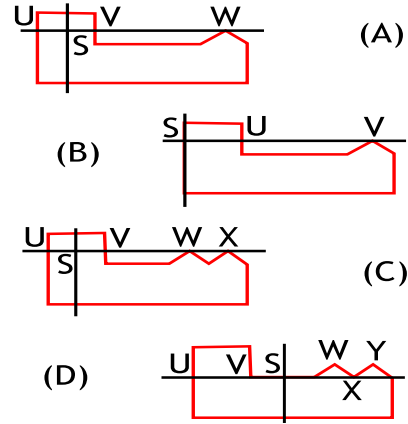


Figure 11: Polygons with locations of the point \mathcal{S} .

Cases for theorem 3 are presented with visual representations in figures 11 and 12. In figure 11.(A), 3 chains exist of which two are valid. The valid chains are (a) UV and (b) WU. The chain VW is invalid. Thus $Inside_{\mathcal{P}}(\mathcal{S}) = 1$ and $Inacc_{\mathcal{P}}(\mathcal{S}) = 1$. In figure 11.(B) the point lies on the edge and the polygon can be divided into 3 chains of which only two contain \mathcal{S} . These valid chains are (a) SU and (b) VS. Thus $Inside_{\mathcal{P}}(\mathcal{S}) = 1$ and $Inacc_{\mathcal{P}}(\mathcal{S}) = 1$. In part (C) of figure 11, the horizontal line cuts through two edges and touches two vertices. Thus there exists 4 chains of which two are valid, namely (a) UV and (b) XU which contain \mathcal{S} . Thus $Inside_{\mathcal{P}}(\mathcal{S}) = 1$ and $Inacc_{\mathcal{P}}(\mathcal{S}) = 1$. Lastly, in part (D) of the same figure, \mathcal{S} lies on an edge and the horizontal line passes through a vertex. In this case, there exist 5 chains of which two contain the point of test and are thus valid. They are (a) VW and (b) YU. Thus $Inside_{\mathcal{P}}(\mathcal{S}) = 1$ and $Inacc_{\mathcal{P}}(\mathcal{S}) = 1$.

Next, in figure 12.(A) 8 chains exist namely, (a) YZ (b) ZW (c) WB (d) BU (e) UV (f) VA (g) AX and (h) XY, of which 6 are valid except UV and XY. These two do not contain \mathcal{S} .

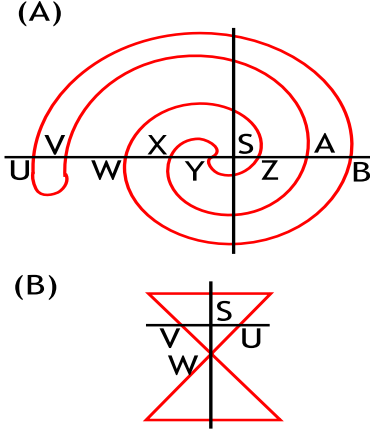


Figure 12: Polygons with locations of the point \mathcal{S} .

Now, the pair that needs to be *broken* is YZ and ZW, while the pairs (WB, AX) and (BU, VA) are to be *ignored* from consideration. Thus, $Inside_{\mathcal{P}}(\mathcal{S}) = 1$ and $Inacc_{\mathcal{P}}(\mathcal{S}) = 1 + 2$ (i.e. one pair requires to be broken and two need to be ignored from consideration). Finally, in figure 12.(B) two chains exist of which both are valid, i.e. (a) UV and (b) VWWU. Thus $Inside_{\mathcal{P}}(\mathcal{S}) = 1$ and $Inacc_{\mathcal{P}}(\mathcal{S}) = 1$.

Theorem 5. *Given a point \mathcal{S} is not inside as well as inaccessible w.r.t polygon \mathcal{P} , then the following statements, (1) $Inside_{\mathcal{P}}(\mathcal{S}) = 0$ and (2) $Inacc_{\mathcal{P}}(\mathcal{S}) = \mathcal{N}$, are logically equivalent: $Inside_{\mathcal{P}}(\mathcal{S}) = 0$ iff $Inacc_{\mathcal{P}}(\mathcal{S}) = \mathcal{N}$*

Proof. (a) If $Inacc_{\mathcal{P}}(\mathcal{S}) = \mathcal{N}$ then $Inside_{\mathcal{P}}(\mathcal{S}) = 0$.

Given $Inacc_{\mathcal{P}}(\mathcal{S}) = \mathcal{N}$ implies \mathcal{N} pairs of valid chains need to be *ignored*. Again, a valid chain by definition is one whose epigraph or hypograph contain \mathcal{S} . If \mathcal{L}_v , i.e. $x = x_0$, is drawn such that it cuts the valid chains and passes through \mathcal{S} , then the chains can be sorted according to the value of intersection

points in $x = x_0$. Let $\mathcal{C}_1, \dots, \mathcal{C}_{2 \times \mathcal{N}}$ be the sorted order of chains from bottom to top. Taking consecutive pairs of these valid chains, i.e. $(\mathcal{C}_1, \mathcal{C}_2), (\mathcal{C}_3, \mathcal{C}_4), \dots, (\mathcal{C}_i, \mathcal{C}_j), \dots, (\mathcal{C}_{2 \times \mathcal{N}-1}, \mathcal{C}_{2 \times \mathcal{N}})$, it is easy to know whether \mathcal{S} is an affine combination of $(x_0, y_{\mathcal{C}_k}^{int})$ and $(x_0, y_{\mathcal{C}_{k+1}}^{int})$ $\forall k \in \{1, 3, \dots, 2 \times \mathcal{N} - 1\}$. Since \mathcal{N} pairs need to be *ignored*, it is evident that \mathcal{S} is not an affine combination of any of the above pairs. This suggests that there does not exist a pair such that $\mathcal{S} \in epi(f_{\mathcal{C}_k})$ and $\mathcal{S} \in hypo(f_{\mathcal{C}_{k+1}})$, that can be *broken*. Since no such pair exists, the status of \mathcal{S} related to \mathcal{P} is $Inside_{\mathcal{P}}(\mathcal{S}) = 0$, which is the desired result.

(b) If $Inside_{\mathcal{P}}(\mathcal{S}) = 0$ then $Inacc_{\mathcal{P}}(\mathcal{S}) = \mathcal{N}$.

Given $Inside_{\mathcal{P}}(\mathcal{S}) = 0$ implies there does not exist a pair of chains \mathcal{C}_i and \mathcal{C}_j such that $\mathcal{S} \in epi(f_{\mathcal{C}_i})$ and $\mathcal{S} \in hypo(f_{\mathcal{C}_j})$. Thus it is difficult to proceed with the proof. Instead, by proving its contrapositive the above statement will hold. Thus if $Inacc_{\mathcal{P}}(\mathcal{S}) \notin \{\mathcal{N}\}$ then $Inside_{\mathcal{P}}(\mathcal{S}) \notin \{0\}$. Since $Inacc_{\mathcal{P}}(\mathcal{S}) \notin \{\mathcal{N}\}$, it implies that $Inacc_{\mathcal{P}}(\mathcal{S}) \in \{1, 1 + \mathcal{N}\}$. It has been proved in part (a) above that if $Inacc_{\mathcal{P}}(\mathcal{S}) \in \{1, 1 + \mathcal{N}\}$ then $Inside_{\mathcal{P}}(\mathcal{S}) \in \{1\}$. But $Inside_{\mathcal{P}}(\mathcal{S}) \in \{1\}$ also means that $Inside_{\mathcal{P}}(\mathcal{S}) \notin \{0\}$. Thus $Inacc_{\mathcal{P}}(\mathcal{S}) \notin \{\mathcal{N}\}$ implies that $Inside_{\mathcal{P}}(\mathcal{S}) \notin \{0\}$. Since the contrapositive holds, so does the original statement. \square

Two cases for theorem 5 are depicted in figure 13. In figure 13.(A) 8 chains exist namely, (a) UV (b) VA (c) AX (d) XY (e) YZ (f) ZW (g) WB and (h) BU. Of these, 3 pairs exist, all of which are valid chains but need to be *ignored* as non of them enclose \mathcal{S} as an affine combination of two intersection points on $x = x_0$. These pairs are (WB, AX), (XY, ZW) and (VA, BU). Thus $Inside_{\mathcal{P}}(\mathcal{S}) = 0$ and $Inacc_{\mathcal{P}}(\mathcal{S}) = 3$. For the case of intersecting polygon in figure 13 (B) 6 chains exist namely, (a) ZU (b) UY (c) YW (d) WX (e) VX and (f) XZ, of which WY and XZ

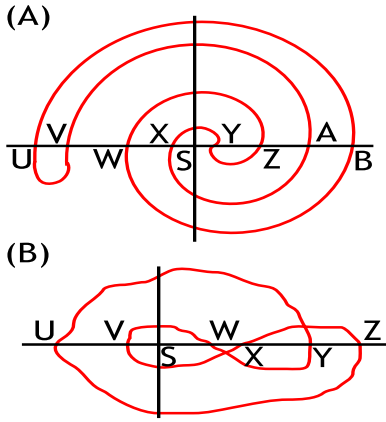


Figure 13: Polygons with \mathcal{S} outside.

are not valid. Remaining pairs of valid chains need to be *ignored* and thus $Inside_{\mathcal{P}}(\mathcal{S}) = 0$ and $Inacc_{\mathcal{P}}(\mathcal{S}) = 2$. For the next few sections, let EH (epi/hypo-graph method) denote the proposed method.

3.3 Crossover vs EH

Crossover (Cr) states that when a line drawn from a point \mathcal{S} in a direction, cuts the polygon \mathcal{P} odd number of times, then \mathcal{S} is inside \mathcal{P} , i.e.

$$Inside_{\mathcal{P}}^{CR}(\mathcal{S}) = \begin{cases} 1, & \text{odd intersections} \\ 0, & \text{even intersections} \end{cases}$$

For the case of line passing through vertices, the problem is solved by shifting the line infinitesimally. Two issues arise in this case: (1) There can be two solutions, if the line is not shifted slightly. (2) If the crossover has to be repeated several times until it finds an odd number of intersections, then it is a nondeterministic problem, in case the line is shot randomly.

By (1) ambiguity arises on the way a ray or line is shot from \mathcal{S} and by (2) nondeterminism arises due to repetition of the procedure of shooting the line randomly. The following figures will illustrate these issues in detail. In contrast to the

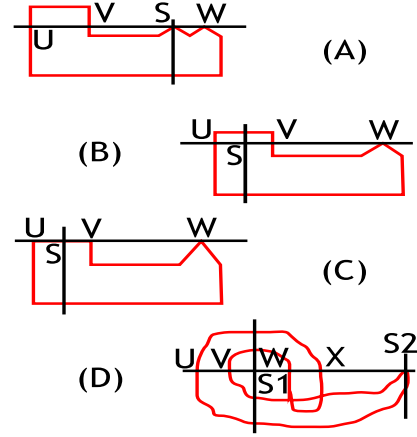


Figure 14: Cases to compare Cross Over (CR) and the proposed method (EH)

Cr, by using theorems 1 and 3, the EH or the proposed method can easily determine if \mathcal{S} lies in \mathcal{P} or not, deterministically. This is because whichever way a line is drawn through \mathcal{S} , if it cuts the polygon, then it will dismember \mathcal{P} into a finite number of countable chains. If it doesn't cut the polygon and \mathcal{S} is a vertex, then also there exist atleast one chain that contains \mathcal{S} . Searching for these valid chains and then locating which of those need to be *broken* is deterministic.

The figure 14 shows the different cases for the comparison of CR and EH method for the same point of investigation. In figure 14.(A), if the a horizontal line is drawn to the left of the \mathcal{S} , then it intersects at two points U and V and if it drawn to the right, it intersects at the point W. According to CR, when the line is drawn to the right of the \mathcal{S} , then \mathcal{S} is inside the polygon. If the line is drawn to the left of \mathcal{S} , then the point is outside the polygon. This is definitely a case of ambiguity. Also, the outcome of the CR depends on the direction of the ray that is shot from \mathcal{S} . This makes the outcome of the test nondeterministic in the sense that it is not known which ray would give the correct result, if the

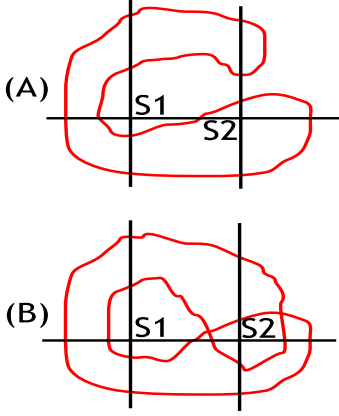


Figure 15: Cases to compare Winding Number Rule (WNR) and the proposed method (EH)

rays are shot randomly.

The EH method overcomes this problem by segmenting the polygon into finitely countable chains. The search for an affine combination of valid chains that may contain \mathcal{S} is deterministic as there are only limited number of chains available for checking. Thus the outcome is singular and deterministic. If two perpendicular rays with its intersection point as \mathcal{S} are drawn at a different orientation, intersecting the polygon at different places, even then by rotating the oriented axis and the polygon to horizontal vertical frame, the solution remains the same. Thus randomness of the directionality of the shot rays do not affect the outcome of the point inclusion test for \mathcal{S} . For part (A) in the figure 14, by CR $Inside_{\mathcal{P}}^{CR}(\mathcal{S}) = (0, 1)$ depending on the number of intersections that is $(2, 1)$. By EH method, $Inside_{\mathcal{P}}^{EH}(\mathcal{S}) = 1$ and $Inacc_{\mathcal{P}}^{EH}(\mathcal{S}) = 3$ by theorem 1. It must be noted that the inaccessibility of the point related to the polygon may change but the status of \mathcal{S} related to \mathcal{P} captured by the definition of *Inside* will not change if the point is inside the polygon.

Similarly, for the part (B) and (C) in figure 14,

by CR the $Inside_{\mathcal{P}}^{CR}(\mathcal{S}) = (1, 0)$ depending on the number of intersections based on the direction of the ray which is $(1, 2)$. Finally, in figure 14.(D), for point \mathcal{S}_1 four valid chains exist namely, (a) VW (b) XU (c) $U\mathcal{S}_2$ and (d) \mathcal{S}_2V , none of which need to be *broken* or *ignored*. Thus by theorem 3 $Inside_{\mathcal{P}}^{EH}(\mathcal{S}_1) = 0$ and $Inacc_{\mathcal{P}}^{EH}(\mathcal{S}_1) = 2$. By CR the outcome of the inclusion test changes, that is $Inside_{\mathcal{P}}^{CR}(\mathcal{S}_1) = (0, 1)$ depending on the intersections obtained by the direction of the ray that is $(2, 3)$. For the point \mathcal{S}_2 , two valid chains exist namely (a) \mathcal{S}_2V and (b) VWXU. \mathcal{S}_2 . Thus by theorem 1, $Inside_{\mathcal{P}}^{EH}(\mathcal{S}_2) = 1$ and $Inacc_{\mathcal{P}}^{EH}(\mathcal{S}_2) = 2$. By CR $Inside_{\mathcal{S}_2}^{EH}(\mathcal{S}_2) = 0$ as the number of intersections is 4.

3.4 Winding Number Rule vs EH

The winding number rule (WNR) states that the number of times one loops round the point \mathcal{S} before reaching the starting point on polygon \mathcal{P} , decides the number of times whether \mathcal{S} lies inside \mathcal{P} or not. Thus

$$Inside_{\mathcal{P}}^{WNR}(\mathcal{S}) = \begin{cases} \mathcal{N}, & \mathcal{N} \text{ loops around } \mathcal{S} \\ 0, & \text{zero loops around } \mathcal{S} \end{cases}$$

An analogy of a prison wall shown in figure 15 is taken into account to explain the differences. Figure 15.(A) is the initial structure of the prison and then the final structure is shown in the part (B) of the same figure. Initially, via the WNR, \mathcal{S}_1 was lying outside and \mathcal{S}_2 inside the prison wall. Same is the verdict by the new method. Next a portion of the prison wall is extended and the final structure looks like that in figure 15.(B). Note that \mathcal{S}_1 and \mathcal{S}_2 are still outside the new prison via the new definition, as the areas in which \mathcal{S}_1 and \mathcal{S}_2 lie, are not reachable from prison's perspective. This is because two pairs of walls have to be *ignored and not broken* in each case. From this point of view both \mathcal{S}_1 and \mathcal{S}_2 are outside \mathcal{P} , in figure 15.(B). Also, even though $WNR = 2$ for \mathcal{S}_2 in new prison in part (B) of the same figure, implying the point lies twice inside, it does not make sense. It can be stated that if a

point lies inside once, then it lies forever. There does not arise the idea of point lying inside \mathcal{N} times. Thus a point lying inside \mathcal{N} times, is the same as point lying inside once. If it does not lie inside, then it won't lie forever. In this way the new definitions and the accompanying theorem are definitive in producing a concrete answer via means of epigraph-hypograph method.

4. TIME COMPLEXITY

Given the amount of computational steps needs to converge to a singular and deterministic solution, it is expected that the algorithm is not fastest. Here the computational time complexity of the algorithm is addressed.

With n edges, the algorithm takes a constant time of C_1 to process the first step mentioned in section 2.1. C_1 is the constant number of steps used in (1) computing the intersection of \mathcal{L}_h with each of the edges and (2) deciding whether the intersection point lies on the edge and between the vertices of the edge. This process takes up $C_1 \times n$ steps. The newly found intersection points are added to the pre-existing vertex list of the polygon while some original vertices close to the intersecting points in the range of $\pm 10^{-5}$ or less are removed to circumvent the error in computation due to floating point precision. This execution step on the whole increases the number of vertices by a fraction, say f_1 were $0 \leq f_1 < 1$. Thus the total number of vertices on the polygon now amount to $(1 + f_1) \times n$.

The decomposition of the polygon into different chains requires the processing of all $(1 + f_1) \times n$ edges, as has been described in section 2.2. This processing is of the order of $(1 + f_1) \times n$ with some constant C_2 required for checking if the chains are valid or not. In section 2.3, as in section 2.1 the points of intersection are computed for \mathcal{L}_v and one of the edges in each of the chains. Since the number of edges belonging to the chains is a fraction f_2 ($0 \leq f_2 < 1$) of the total number of edges of the polygon, the execution of this step requires $C_1 \times f_2 \times (1 + f_1) \times n$ units of cpu time.

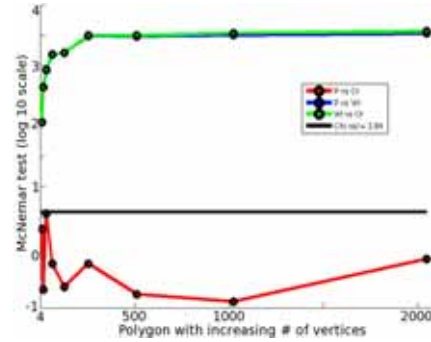


Figure 16: Check of statistical significance using McNemar's Test. Significance level: $\alpha = 0.05$. Critical χ^2 value of 3.84 $\Rightarrow p < 0.95$ with 1 degree of freedom. Values higher than the critical value imply significant differences with $p \geq 0.95$. Black horizontal line $y = 3.85$ with a p value of 0.95.

Finally, after sorting the chains (which requires $m \times \log(m)$ where $m = f_2 \times (1 + f_1) \times n$), checking whether the sample point is an affine combination, requires a constant time of C_3 . This procedure is mentioned in the previous section. Summing up the total time of execution, the algorithm works in $(C_1 \times n + (1 + f_1) \times n + C_2 + C_1 \times f_2 \times (1 + f_1)n) + C_3 + m \log(m) = (C_1 + 1 + f_1 + C_1 \times f_2 + C_1 \times f_2 \times f_1) \times n + C_2 + C_3 + m \times \log(m)$. Let $A = (C_1 + 1 + f_1 + C_1 \times f_2 + C_1 \times f_2 \times f_1)$ and $B = C_2 + C_3$, then the computational time complexity of the algorithm is $A \times n + B + m \log(m) \approx \mathcal{O}(n \log n)$. This time in computation is the worst case scenario, where all the edges have to be processed in for all the four main steps described above.

5. RESULTS AND DISCUSSION

The main aim of this manuscript is to theoretically analyze the correctness of the proposed method. It would be interesting to know how algorithms based on the two widely used cross-over and winding number rule concepts fair, apropos to the algorithm based on the proposed method.

The tests in this manuscript are by no means complete. Widely used implementations in matlab have been adopted for comparison purpose. The aim of these tests is to study the overall accuracies of cross-over and winding number rules with respect to the proposed method. The authors invite the community of computational geometry to put the foregoing method to rigorous empirical test.

One of the measures of fairness is the test of significance of results obtained from the algorithms. McNemar's Test [15] was employed to check the statistical significance of one algorithm against another. In a 2×2 contingency table, the McNemar's test gives a statistic similar to the chi squared statistic which is formulated as:

$$\chi^2 = \frac{(|n_{01} - n_{10}| - 1)^2}{n_{01} + n_{10}} \quad (3)$$

where, n_{01} and n_{10} are the false positives and the false negatives, respectively. Equation 3 was used as it contains the correction for discontinuity. In this study using McNemar's test, the significance level $\alpha = 0.05$ and χ^2 value of 3.84 $\implies p < 0.95$ with 1 degree of freedom. Values higher than the critical value imply significant differences with $p \geq 0.95$.

Artificial polygons with number of vertices in the set 2^i where $i \in \{2, \dots, 11\}$ were generated. For a particular number of vertices, 10000 polygons were generated and stored. Each polygon was generated by randomly generating the co-ordinates and joining the vertices in order. The last vertex was finally joined with the first vertex. This created a series of intersecting as well as non intersecting polygons. All polygons were generated in the bounding box of $[0, 1]$. For each of the 10000 polygons having a particular number of vertices, a random test point was generated in the bounding box.

Let (P) be the proposed algorithm, (Cr) the Crossing over algorithm and (Wi) the Winding number rule algorithm. A Matlab version of the crossing over algorithm was adopted from [16]. Matlab's [17] inpolygon algorithm was adopted,

W.r.t P	Error		Fpr		F-Score	
No. vertices	Cr	Wi	Cr	Wi	Cr	Wi
4	0	0	0	0	1.00	1.00
8	0.0002	0.01	0.0003	0.01	0.99	0.97
16	0.0005	0.04	0.0003	0.06	0.99	0.93
32	0.0010	0.08	0.0003	0.12	0.99	0.88
64	0.0017	0.15	0.0012	0.25	0.99	0.83
128	0.0018	0.16	0.0016	0.26	0.99	0.82
256	0.0186	0.32	0.01	0.58	0.97	0.73
512	0.0485	0.32	0.04	0.56	0.94	0.73
1024	0.1339	0.36	0.12	0.66	0.85	0.71
2048	0.2373	0.39	0.22	0.73	0.75	0.70

Table 1: Error, Fpr and F-score for 10,000 samples for every number of vertex, obtained for Cr and Wi w.r.t P (the proposed algorithm).

which is an implementation of the winding number rule algorithm. The proposed algorithm has been coded in Matlab [17] and made available. Note that the paper provides new definitions and resolves some fundamental subtle issues raised by the Cr and Wi definitions. It does not by any means claim that the implemented code is the best in comparison to other existing Matlab codes for Cr and Wi.

Results in figure 16 show the statistical significance in the performance of the algorithms against each other. Clearly, it can be seen that the results generated by Cr is significantly similar to those generated via P (red graph in the figure). But this may not be true for polygons with very large vertices (which needs to be checked) as the red curve begins to move upwards towards more significant region (above the horizontal line in the aforementioned figure).

While comparing the results of Wi with those of Cr as well as P, it can be seen that the results are significantly different as the number of vertices increases (blue and green graphs in the figure). This apparent statistical difference is due to the fact that if the winding number is $\ell \neq 0$, then the point lies ℓ times inside the polygon. This is not true with the Cr which states that even

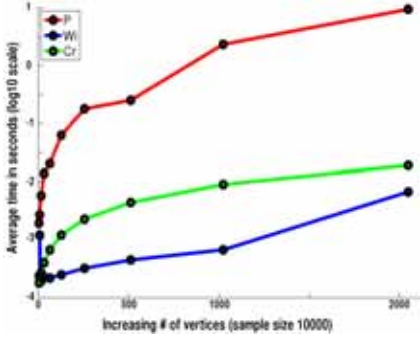


Figure 17: Average time in seconds on a \log_{10} scale for polygons with increasing number of vertices. Sample size of 10000 for each number of vertices.

cross overs imply the that points are outside the polygons. P, on the other hand sets a bound on the winding number rule, stating that the if $\ell \neq 0$, then the point lies inside the polygon if and only if the criterion of having a pair of chains (one above and other below the point) is true (via proved theorems above). Thus even if the winding number evaluates to say 2, in case of the star polygon for the point in figure 4, P will always give the result that the point lies outside the polygon and not twice inside the polygon.

Table 1 shows the accuracy of the results obtained by the Cr and Wi, with respect to the assumed ground truth generated by the proposed algorithm P. Increasing rate of error of Wi in comparison to Cr, strengthens the explanation in the foregoing paragraph. Even though Cr has been shown not to be statistically significant from P, with increasing number of vertices (sample size of 10000), the error rate creeps up. The accuracy by itself does not always suffice to give the measure of correctness and thus is aided via means of the false positive rate, precision and recall. The false positive rate in table 1 tells how much one algorithm generated falsely true results given that base algorithm P had labeled the results as false. The results suggest that both the Cr and Wi start

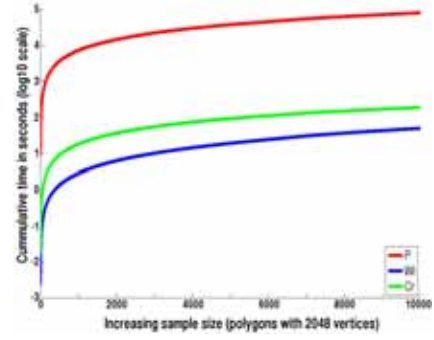


Figure 18: Cumulative time in seconds on a \log_{10} scale for increasing sample size for polygons with number of vertices 2048.

contributing to the false positives with respect to P as the number of vertices increases. Wi gives greater false positives with respect to Cr. Again, this can be attributed to the above explanation in previous paragraph.

The F-score measures in table 1 suggest the quality of the concentration or precision and the retrieval or recall of Cr and Wi with respect to P. It can be seen that the F-score for Cr remains much closer to 1 than Wi, which falls drastically with increasing number of vertices. The score indicates proximity of the results obtained by the proposed algorithm w.r.t Cr and Wi. For polygons with small number of vertices the Cr follows close to P, but this pattern does not remain the same for large number of vertices. This can be attributed to the rare instances that the proposed algorithm is able to solve where even Cr may fail.

Finally, the average time in seconds on a log scale has been plotted in figure 17 to depict the time consumption of P, Cr and Wi. It is found that due to large computation requirement, the time consumption (t) takes the form $t_P > t_{Cr} > t_{Wi}$. Also, in figure 18, for polygons with the number of the vertices 2048, and an increasing sample size, the cumulative time in seconds is plotted on the log scale. The graph suggest that the reliability of the proposed solution comes at a com-

putational cost. Apart from that, the correctness of the proposed solution is not affected, which forms the kernel of the manuscript.

6. CONCLUSION

A theoretically reliable and an analytically correct solution to the point in polygon problem is proposed, using epi-hypo graphs, homotopy and Hopf's Degree Theorem. The proposed definitions along with the partial proofs indicate the bridge between the cross-over and the winding number rules. The method also resolves the problem of ambiguity of solution due to the ray passing through vertices, the direction of the shift of the ray, and nondeterminism induced by usage of the cross over multiple number of times. For complex self-intersecting polygons, the presented algorithm challenges and addresses the issue of point being multi-ply inside the polygons posed by the winding number rule. By providing an unambiguous, singular and deterministic solution for both non-intersecting and self-intersecting polygons, the novel solution redefines concepts of inside and inaccessibility of a sample point with respect to a polygon and presents a fresh perspective of solving the old point in polygon problem.

7. ACKNOWLEDGEMENT

The authors thank the NeuroImaging center at UMCG Groningen, in The Netherlands, for supporting the project as a part of research in development of brain image segmentation algorithms using computational geometry algorithms.

REFERENCES

- [1] B. Zalik and I. Kolingerova. A cell-based point-in-polygon algorithm suitable for large sets of points. *Computers and Geosciences*, 27:10, 1135-1145, 2001.
- [2] M. Chen, and P. Townsend. Efficient and consistent algorithms for determining the containment of points in polygons and polyhedra. *Eurographics*, 423-437, 1987.
- [3] C.W. Huang and T.Y. Shih. On the complexity of point-in-polygon algorithms. *Computers and Geosciences*, 23:1, 109-118, 1997.
- [4] S. Nordbeck and B. Rystedt. Computer cartography point in polygons program. *National Swedish Institute for Building Research Report*, 19 pages, 1967.
- [5] U. Manber. *Introduction to algorithms - a creative approach*, Addison-Wesley, 478 pages, 1989.
- [6] J.D. Foley, V. Dam, A. Feiner and J.F. Hughes. *Computer Graphics - Principles and Practice*, Addison-Wesley, 1174 pages, 1990.
- [7] E. Haines. Point in Polygon. *Strategies. Graphics Gems IV, chapter 1.4*, 24-46, 1994.
- [8] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.
- [9] R.V. Eynde. Historical evolution of the concept of homotopic paths. *Archive for History of Exact Sciences, Springer*, 45:2, 127-188, 1992.
- [10] D. Siersma. Poincaré and Analysis Situs, the beginning of algebraic topology. *Nieuw Archief voor Wiskunde. Ser. 5, Koninklijk Wiskundig Genootschap*, 13:3, 196-200, 2012.
- [11] J.M. Moller. Homotopy theory for beginners. <http://www.math.ku.dk/~moller/e01/algtopI/comments.pdf>, 2012.
- [12] H. Poincaré. Papers on Topology: Analysis Situs and Its Five Supplements. *American Mathematical Society*, 37, 2010.
- [13] T. Needham. *Visual complex analysis*. Oxford University Press 1999.

- [14] J. Bosshardt. The hopf degree theorem. <http://www.math.uchicago.edu/~may/VIGRE/VIGRE2011/REUPapers/Bosshardt.pdf>, 2011.
- [15] Q. McNemar. Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika*, 12, 153-157, 1947.
- [16] D. Engwirda. Inpoly function, <http://www.mathworks.com/matlabcentral/fileexchange/10391>, 2007.
- [17] Matlab inpolygon function. <http://www.mathworks.com/access/helpdesk/help/techdoc/ref/inpolygon.html>
- [18] M. Shmrat. Algorithm 112: Position of point relative to polygon. *Communications of the ACM*, 5:8, page 434, 1962.
- [19] A.S. Glassner. An introduction to ray tracing. *Morgan Kaufmann*, 1989.
- [20] F.P. Preparata and M.I. Shamos. Computational Geometry. *Springer*, 1-35, 1985.
- 3. Remco J Renken works as MR Physicist at the Neuro Imaging Center in University Medical Center Groningen, The Netherlands. His interests include working on fMRI data analysis and conducting fMRI experiments.
- 4. Gert J ter Horst is professor and General Director of the Neuro Imaging Center in University Medical Center Groningen, The Netherlands. His interest is in various problems involving neuro scientific research.

ABOUT THE AUTHORS

- 1. Shriprakash Sinha worked as a Researcher at the Neuro Imaging Center in University Medical Center Groningen, The Netherlands. His interest include image segmentation, computational geometry and applications of machine learning algorithm to real life problems.
- 2. Luca Nanetti worked as a Researcher at the Neuro Imaging Center in University Medical Center Groningen, The Netherlands. His interests include fMRI data analysis and connectivity analysis via DTI in neuro-scientific problems.

QUALITATIVE INVESTIGATION ON FORMATIVE DESIGN PROCESS USING 3D CAD BY INDUSTRIAL DESIGNERS

Misako NISHII and Takafumi SAITO

Tokyo University of Agriculture and Technology

ABSTRACT: On 3D CAD operation by industrial designers, their operation skill levels and characteristics of CAD systems affect quality of created shapes. The final aim of this study is to find out new technical functions of 3D CAD which minimizes such skill and system dependency. In this present study, we observed and analyzed 3D modeling process of designers. In a process of 3D formative design, each designer operated 3D CAD system until he/she has obtained the basic curves which constitute the intended shape. Designer with skillful operation techniques in 3D CAD were selected as the subject of investigation, because it is easy to pick up objective opinions of operation. As the target object to design, we chose a faucet (a cock of water supply) to be installed in public spaces, since it is easy to understand designer's intention. As a result of investigation, we found the following: 1) sensuous impressions internalized in design such as "heavy", "hard" and "rugged look" are visualized as curves from the initial stage of design process; 2) designer edits the shape by fine-tuning the curves so that the impressions are visualized into the intended shape, while designer is keeping the whole proportion of the shape in the process to completion; 3) the number and positions of control points and degree of each curve depend on designers' learning method under certain working environments such as advanced design or mass-production system. We observed that designer cannot always control their curves based on complete understanding of geometric features of the curve.

Keywords: 3D CAD, Modeling, Formative design.

1. INTRODUCTION

This study focuses on 3D CAD systems for 3D formative design. Designers use such these tools by themselves to create 3D shapes on computers. Here, we focus on free-form curves used in 3D CAD.

We infer that in order to accurately reflect the intention on 3D CAD, a designer is required to obtain three kinds of ability as follows:

- 1) Observation to refine whether intended shapes are obtained;
- 2) Judgment of extent of modified shape;
- 3) Skillful operation capability on 3D CAD.

It is necessary for designers to continue long-term effective practical experience to

master these kinds of ability, which generates synergistic effect through association of sensibility and skill.

In the three kinds of ability, "skillful operation capability" is affected by practical experience and quality of training operations. Whether or not a designer can obtain an intended shape on 3D CAD depends on designer's operation skill. It takes an enormous amount of time to learn and obtain a satisfactory level of operation skill. For that reason, it is necessary to improve 3D CAD systems more useful and convenient for designers.

In this study, we observe and analyze 3D modeling process of designers. In a process of 3D formative design, each designer operated

3D CAD system until he/she has obtained the basic curves which constitute the intended shape.

2. BACKGROUND OF THIS STUDY

The needs of this study are presented here.

2.1 Transition of 3D CAD in product development site

3D CAD has been developed as a system to aid engineering, and it has been continuously improved. In recent years, ICT (Information Communication Technology) and 3D CAD have been introduced into workplaces all over the world in order to improve products. It is one of the communication tools for designers to understand 3D designed shape data. And it has been used in the field of design workplace for embodying the idea of designers.

Currently, input methods of 3D shape used in design workplace can be classified into 4 categories.

- 1) 3D scanning of mock-up (prototype of the external design) [1]
- 2) Manual 3D design on CAD system based on 2D hand-drawn design sketch [2]
- 3) Automatic generation of 3D data from 2D hand-drawn design sketch
- 4) Direct 3D design on CAD screen with a pen-type input device [3]

As previous studies, for example, Harada *et al.* showed a set of systematized aesthetic curves with characteristics of aesthetic curves in natural and artificial real objects, which can be used as hints for designers [4]. They proposed an automatic generation method of aesthetic spline curves from hand-drawn sketch curves [5].

However, even if these improvements have been made, the present situation still stays that designers require hard work to learn 3D CAD operation and 3D modeling. Considering on that premise, the present 3D CAD is not useful enough for designers to use in subjective sensation.

2.2 Fragment of 3D CAD education in design

field

We focused on CAD system manuals, which are one of the learning operations of 3D CAD for designers.

A Book "Rhinceros ver.5 NYUMON" [6] is a beginner's operation manual of NURBS modeling tool. The author is a designer who has used the 3D CAD for more than 10 years. He has experienced to try out various 3D modeling tools and wrote the articles tips based on those experiences. He is in charge of the lecture for 3D CAD training in industrial design.

One example of the operation procedures in this book is shown below. It is about the operation when drawing the curve.

STEP1: Execute "curve" command.

STEP2: Specify "degree 3".

STEP3: Put control points at predetermined positions.

However, the following reasons have not been described: why should the degree be 3?; why should the control points be put at the position? We have asked the questions to the author, and he answered he did not intentionally avoid the description about the degree, because it was so difficult for beginners to understand.

The "Rhinceros DE KIWAMERU 3D digital design" [7] is another Rhinceros operation manual written by another author. This book described degree, knot and weight in this book. It means designers must understand NURBS structure mathematically. We wonder how long it takes until they obtain the way to construct optimum curves for intended shapes and structured data. We also wonder how the designers who began to use 3D CAD in early days managed to obtain or master the ability by their independent learning process and efforts. To begin with, do designers master ability for 3D CAD perfectly?

2.3 Point of view on digital fabrication

Recently, digital fabrication has been attracting attention in manufacturing [8]. The advent of the 3D printer has dramatically changed the task of making the prototype mod-

el. The manufacturing environment is a major difference to the conventional manufacturing method. Finished products can be manufactured without using a mold. Further, it is not necessary to consider the characteristics of the cutting and draft angles. Using 3D printers allow designers to manufacture finished products without having the conventional manufacturing environment. As a result, without having the conventional manufacturing environment any designer and common users can manufacture original products. At such occasions, user friendly 3D modeling tools are critical.

2.4 The point of view of Kansei engineering and Cognitive science

We focus on "Observation to refine whether the intended shapes are obtained" which was described in Chapter 1. Cognitive science and Kansei engineering seem preferable to study and solve impressions quantitatively. In order to help Cognitive science and Kansei engineering with the study, designer has to select better 3D CAD first of all.

Therefore, these research fields were excluded from the target at this timing.

3. INVESTIGATION

We focused on 3D molding by skillful designers, and we executed qualitative investigation of their modeling process. We thought that we could pursue critical and necessary functions for the designers with a thought of picking up and considering lacking functions in the present 3D CAD. We selected four designers for this investigation. The rationale for the object or selection standards are described in the following.

3.1 Selection of designers

Because of the ease to obtain objective opinions about the intent of the operation to use 3D CAD, the subjects were designers applicable to the following criteria (Table 1).

Table 1: Selection criteria item of the subject.

	Items
1	Designers are engaged in the process level concept design planning.
2	Designers have experienced the design business for more than 10 years.
3	Designers have shaped the design shape using 3D CAD at long period.
4	Designers have learned how to resolve the operation until the shape of intended on 3D CAD.
5	3D CAD used here is selected on condition that 3D CAD must be easy to control free curves by the designer himself.

3.1.1 The reason of selected: Criteria item 1

The design works cover a wide range including from planning to design combined with engineering. (see Figure 1).

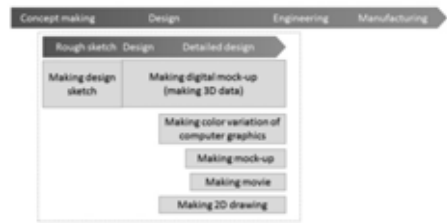


Figure 1: The working process in product development

Regarding the range of design works, all the designers do not get involved in the whole range through. Each designer has his own specified part in the whole range. Interviews were conducted to four designers. This interview allows us to determine what part in the whole range as investigation object is valid in design works. Designers responded about products handled in daily works, the range of design works, and the design procedure using the 3D CAD.

In the first stage of the procedure, Phase 1, designers decide concept design of commensurate with the target (sense of the world). Here, they do not create a specific shape.

In Phase 2, designers create a variety of de-

sign shapes on the basis of the design concept planned in Phase 1 (Figure 2). Designers start using 3D CAD from this stage. However, they do not always use 3D CAD.

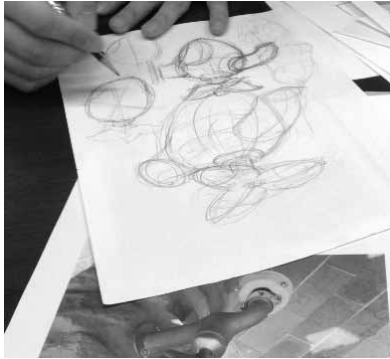


Figure 2: Phase 2

In the next stage, Phase 3, designers continue to pursue the design shape which was created in phase 2. Designers, even though they organize the shape by hand-drawn sketch, the sketch is not a complete shape as design shape.

In Phase 4, designers repeatedly perform the task of shape modification to the design review on the basis of the concept.

Within the range of the process, designers do not design specific shapes in concept stage, and we found that it is not a timing to use 3D CAD. In the design stage close to mass production engineering, designers consider matching with the parts mounted inside. Therefore, in this stage, it is better for the designer to use control through numerical parameters than to use free-form curves.

As a result, we select the designers who engaged in close to the concept design process in the range of the design works as the subjects of our investigation.

3.2 Target object to design

As the target object to design, we chose a faucet to be installed in public spaces. The reason of this selection is that the intended use of the target is simple. Because complex mechan-

ical component is not mounted in the interior of the faucet, it is possible for designer to consciously focus on his modeling.

3.3 3D CAD system

In our investigation, each designer can use the 3D CAD system which he is usually using, because it is implied that the process constituting shape by designers with trial and error are almost the same, even if 3D modeling tools used are different from each other. We carried out investigations for two designers. One is using Rhinoceros ver.5.0 [2], and the other is using SolidWorks2013 [9].

3.4 Details of design process

We had interviews with designers in the usual flow where the process similar to the design process is running usually (Table 2), because we can grasp rationale more deeply from the interviews in each operation, resulting in selecting the best 3D CAD for investigation.

The information we want to get inevitably is about the 3D modeling operation on CAD system based on concept. Accordingly we investigated the curve creation process of the designers through the interview with the designer regarding the designer's concept.

Table 2: Work process of modeling from design to planning.

Process	Work process
1	Designer makes a design concept for the faucet.
2	Designer draws a design shape by hand.
3	Designer creates curves on 3D CAD by referring to the hand-drawn design shape.
4	Designer continues editing the curves until the exact intended shape is obtained on 3D CAD.

3.4.1 Work content in Process 1

Each designer makes design concept. Figures 3 and 4 are examples. This timing is the visualization of the various design approaches.

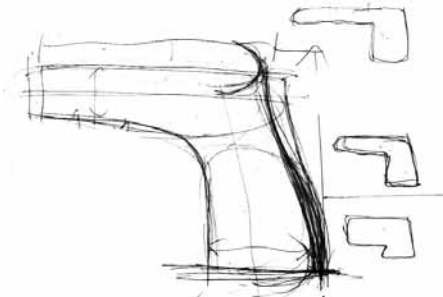


Figure 3: Design sketch by subject A.



Figure 4: Design sketch by subject B.

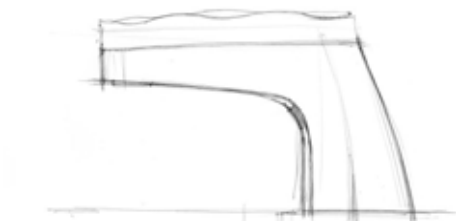


Figure 5: Outline of the side-view.
It can be referred to for 3D modeling.

3.4.2 Work content in Process 2

The designer draws an outline on a paper, which will be used in 3D modeling as a reference. Figure 5 shows an example.

3.4.3 Work content in Process 3

Outline drawing is scanned and input into the 3D CAD in Process 3. Then, the designer continues to perform modeling referring to the outline drawing (see Figure 6).

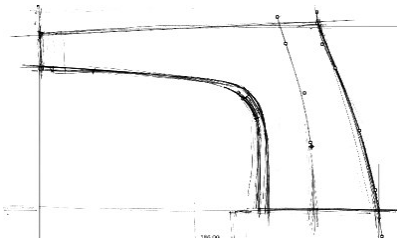


Figure 6: Created curves and hand-drawn design shape

3.4.4 Work content in Process 4

The designer constructs overall shape with curves. He edits the curve shapes repeatedly until he gets satisfactory performance. He fine-tunes while watching the balance of proportion until he obtains an intended shape. Figures 7 and 8 are example results of this process.



Figure 7: Result design of Process 4.

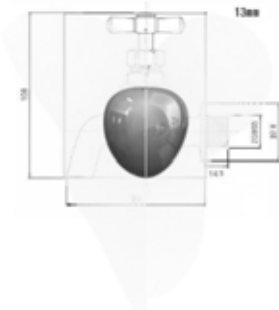


Figure 8: Result design of Process 4.

4. RESULTS AND DISCUSSIONS

In this section, we describe the results of the investigation, especially for curve operations on CAD systems in the Process 3 and 4. Taking a broad view of design process, the curve constitution process does not depend on designers. In the case of SolidWorks, however, we found out that the designer cannot perfectly control the curves. This is because the design is affected by a certain modeling constraint which is a unique feature of the software. Therefore, we mainly describe an investigation of the design process using Rhinoceros.

We added up the curve modeling steps chronologically. Specifically, 3D CAD operations were captured by video, and operations on each curve (A to E shown in Figure 9) constituting the side view shape were classified into 6 categories: "Sketch import", "Curve creation", "Curve editing", "Curve reconstruction", "Camera operation", and "Thinking without CAD operation". Especially, we have summarized about 3 categories concerning curves. From the timing that the designer create or modified each curves, we extracted the amount of time spending on editing and frequency for each curve.

Figure 10 shows the results, where the operation categories for each curve are visualized chronologically to represent the entire operations objectively. In addition, we have enumerated intensions of the operation. Furthermore, we classified each operation whether it comes

from design shape or it is performed in order to increase the accuracy of curve structure.

In our summary of the result, Curve E was excluded, because the bottom portion is a straight line and has little influence.

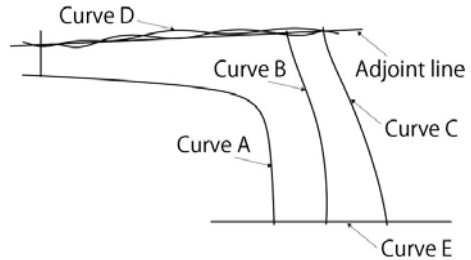


Figure 9: 5 curves in the side view.

4.1 Side shape: the entire operation

Table 3 shows the summary of the 3 categories concerning each curve. Time spent on curve editing and reconstruction in order to obtain the intended shape is 82% of the total curve operations (Figure 11).

A designer used degree 3 when he created a new shape. One reason to use degree 3 is that it is the default value and there is no special reason to draw a curve by changing the degree at this timing. In the editing process of the shape, at the first time, a designer has performed the operation of adjustment of the control point (Figure 12). It is designer's preparation rather than to obtain the intended shape at this timing, and he simply moved the control points for easy editing shape as intended on the later operation. After adjustment, he moved to the next operation. Then he moved the control point for the purpose of obtaining his intended shape and edited curve shape and finally completed the shape (Figure 13).

Table 3: Summary of operation time for the curve construction

	Curve A		Curve B		Curve C		Adjoint line		Curve D		Curve E		Total	Ratio
Curve Creation	16"	8.29%	19"	11.88%	14"	10.77%	7"	0.64%	44"	35.77%	18"	42.86%	118"	18.02%
Curve Editing	126"	65.28%	129"	80.63%	112"	86.15%	0"	0.00%	48"	39.02%	24"	57.14%	439"	67.02%
Curve Reconstruction	51"	26.42%	12"	7.50%	4"	3.08%	0"	0.00%	31"	25.20%	0"	0.00%	98"	14.96%
Total	193"		160"		130"				123"		42"		655"	

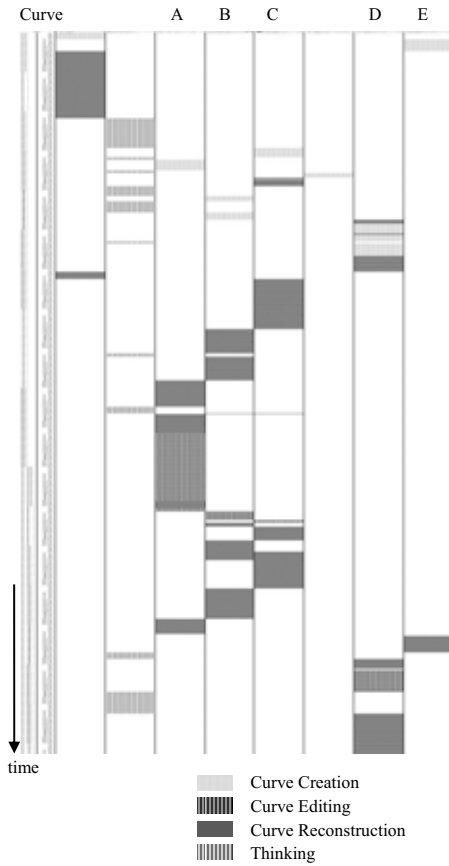


Figure 10: Chronological visualization of curve operations.

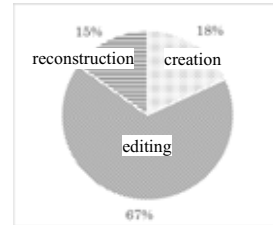


Figure 11: Ratio of each operation in the total works

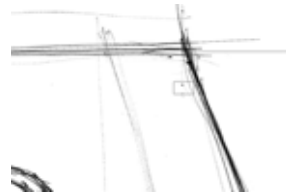


Figure 12: Editing Curve C

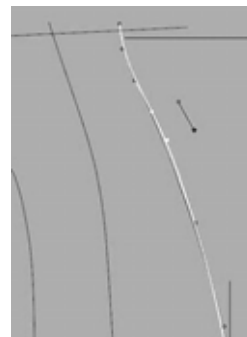


Figure 13: Editing Curve C

4.2 Side shape: Curve A

In Curve A, time spent on creating a new curve was 8.29% of the total. Time spent on editing the curve was 65.28% of the total. Notably, 26.42% of the total was spent on reconstruction of the curve (Figure 14).

The designer cannot be satisfied with a shape even after editing, and tried to rebuild the shape using it. In the command "Rebuild", he increased the number of control points and change the degree from 3 to 5 (Figure 15).

But the designer still cannot be convinced, and had to create a new curve tracing its shape using the "Curve" command. In the case of Curve A, designer had to focusing on where the curvature changes abruptly. A designer had repeated editing many times because he could not complete his intended shape in curved portions. It is revealed that even experienced designer had difficulty in controlling curves which have abrupt curvature.

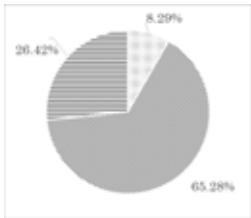


Figure 14: Ratio of each operation in Curve A

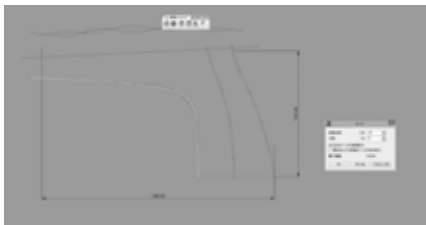


Figure 15: Editing Curve A

4.3 Side shape: Curves B and C

In Curve B, time spent on creating a new curve was 11.88% of the total, and editing the curve was 80.63% of the total (Figure 16 left).

In Curve C, time spent creating a new curve was 10.77% of the total, and editing the curve was 86.15% of the total (Figure 16 right).

A designer has recreated the curve with Rebuild command, like Curve A, but did not use Curve command.

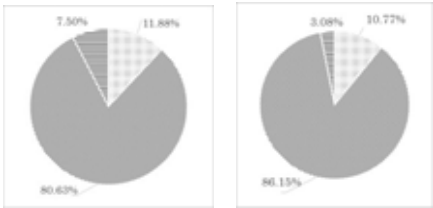


Figure 16: Ratio of each operation in Curves B (left) and C (right)

4.4 Side shape: Curve D

In Curve D, time spent to create a new curve was 35.77% of the total, and editing the curve was 39.02% of the total (Figure 17).

The reason why the ratio of editing is less than other curves is that they have a plan to edit Curve D again after the side shape is arranged. Therefore, the shape is still unfinished at this timing.

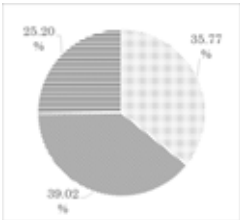


Figure 17: Ratio of each operation in Curve D

4.5 Bottom shape: the entire operation

In the case of the working process of the bottom shape, the idea of designer is to draw one continuous curvature curve which is the combination of almost linear part and elliptic arc part (Figure 18).

To embody this shape, in the first step, he created a part of an ellipse and its tangent lines.

In the next step, he reconstructed a degree 5 curve with Rebuild command. The designer has believed that it is the most efficient way to draw the intended shape. The combination of elliptic arc (degree 2 NURBS) and the tangent lines (degree 1 NURBS) is not the intended shape for designer. It is only utilized as a reference shape.

However, the reconstructed curve with Rebuild command is calculated based on the curve specified by the user, with specified number of control points, within a tolerance. That means Rebuild command function is only to approximate the curves. The meaning of the original curve is not considered at all.

As a result, in this case, the control points of the rebuild curve waved (Figure 19), which was contrary to the intention at all for the designer. He carried out helplessly working to align on the same axis as a workaround for this shape (Figure 20).

In addition, a curve constructed by Rebuild command (Figure 19) has the property that the distance of adjacent two control points near the endpoints becomes shorter. This is natural because the internal Bezier segments which form the entire curve become almost equal length. For designers, however, it is not beautiful visually. He feels that control points locate inappropriate positions.

A beautiful arrangement of control points for the designer means that they should locate apart each other for lower curvature regions, and close each other for high curvature regions, which is his own rules.



Figure 18: Bottom shape: initial input.

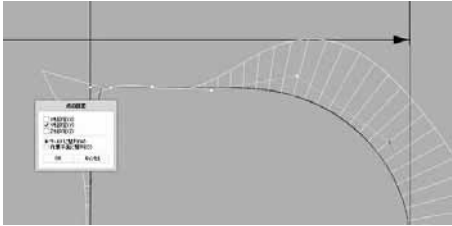


Figure 19: Bottom shape: rebuild.

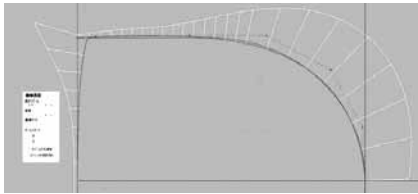


Figure 20: Bottom shape: rearranged control points

5. CONCLUSION AND FUTURE STUDIES

The sensuous impressions internalized in design such as "heavy", "hard" and "rugged look" are visualized as curves from the initial stage of the design sketch. Designers edit the shape by fine-tuning the curves so that the impressions can be visualized into the intended shape, while designers are keeping the whole proportion of the shape in the process to completion. The number and positions of points and degree composing curves depend on designers' learning method under certain working environments such as advanced design and mass-production system.

From our investigation results, it is necessary to analyze the relationship between the intended shape and the optimum curve data structure. In the future, we will continue to carry out the investigation in the same way in order to mount better functions on 3D CAD for designers. In addition, we plan to study beginners and we can survey whether we will have the same results with those of the experienced designers. We believe we can offer the developers clear definite information.

REFERENCES

- [1] The COMET L3D, Steinbichler,
Retrieved Jun 10, 2014,
<http://www.steinbichler.com/products/optical-blue-light-3d-scanning/comet-l3d.html>
- [2] Rhinoceros, Robert McNeel & Associates,
Retrieved Jun 10, 2014,
<http://www.en.na.mcneel.com/>
- [3] Geomagic Freeform, Geomagic,
Retrieved Jun 10, 2014,
<http://geomagic.com/en/products/freeform/overview>
- [4] Y. Harada, and F. Yoshimoto, Analysis of the Curves of Japanese Swords, *Japanese Society for the Science of Design*, (2004) , page 77-84
- [5] K. Masayuki and Y. Harada, Development and application of a system generating aesthetic curves from sketches using visual languages, *IPSJ SIG Technical Report* (2007), page 43-47
- [6] Yasuhisa KOREEDA, *Rhinoceros ver.5 NYUMON*, Rutles, Inc., (2013), pp 21-23
- [7] Atsuo NAKAJIMA, *Rhinoceros DE KI-WAMERU 3D digital design – ver.5 NI KANZEN TAO* -, Rutles Inc., (2013), pp 67-68, 252-254, 264-267
- [8] Ministry of Economy, Trade and Industry Manufacturing Industries Bureau Office of Director for Manufacturing Industries Policy, *SHIN MONODUKURI HOUKOKUSYO*, Retrieved April 10, 2014,
http://www.meti.go.jp/committee/kenkyukai/seisan/new_mono/pdf/report01_02.pdf
- [9] SolidWorks, Dassault Systèmes, SolidWorks Corp., Retrieved Jun 10, 2014,
<http://www.solidworks.co.jp/>

ABOUT THE AUTHORS

1. Misako NISHII is a doctoral program in Graduate School of Bio-Applications and Systems Engineering, Institute of Engineering at Tokyo University of Agriculture and Technol-

ogy. Her research interests shape processing for Designer. She can be reached by e-mail: 50013701291@st.tuat.ac.jp

2. Takafumi SAITO, Ph.D., is a Professor in Institute of Engineering at Tokyo University of Agriculture and Technology. His research interests include computer graphics, visualization, and shape processing. He can be reached by e-mail: txsaito@cc.tuat.ac.jp, or through postal address: Tokyo University of Agriculture and Technology, BASE, 2-24-16, Naka-cho, Koganei, 184-8588 JAPAN.

REAL LIFE AND REAL REPRESENTATION – EDUCATING ALONG AND ACROSS THE BOUNDARIES OF GEOMETRY AND GRAPHICS

Niels-Christian FRITSCHÉ

Technische Universität Dresden, Germany

ABSTRACT: This paper suggests exploring exemplary didactic approaches in analog and digital fashion, as well as 2D versus 3D, a ping pong of geometry, pictures, images and models, with the chance to experimentally cross over from one realm to the other. The physical goal is a portfolio of what I call “direct graphic design”, i. e. graphic design that is produced immediately with the work in front of the producer, where action is constantly echoed by the immediate response of the work, and with little or no guidance by the established rules of the trade (they are taught in other classes). The crossover of categories is encouraged for a number of reasons: First of all speed. The course shall inspire students to work in series to allow comparison and approximation. Second: The large variety of techniques and software branches. Productive mistakes and innovation often occur on the edges of and between fields. Third: The cognitive training that comes with the *mêlée* between 2D and 3D is what characterizes the work of architects. Fourth: The start of a personal library of graphic patterns that are not tied to specific tasks. Fifth: Rules are a powerful tool but not the only guarantee of success. In order to proceed we have to learn the trade to put it aside afterwards. Sixth: Studying architecture all too often appears to be a sequence of unrelated fields of study. We want to bring the opposites together in order to promote integrative thinking.

Keywords: Education, experiment, didactics, artistic fundamentals, geometry and graphics.

1. INTRODUCTION

After the establishment of the education of architects 200 years ago and a first evolution by the Bauhaus in the 1920s the digital revolution could have triggered the discussion about a third era of architectural education. Instead, the debate was split between administrative concerns labeled as “Bologna Process” and legal concerns developed by the various chambers of architects.

Meanwhile, the digital revolution in architecture urged a new parallelism of the acquisition of software skills, the virtualization of architectural design and a stylistic “anything goes”.

A contemporary didactic concept for the education of the architect along the boundaries of geometry and graphics should be based on broad fundamentals. The specifications for these fundamentals sound universal: We want

to deal with the real carbon world because that is where we exist and what we want to work for as architects. We want to anticipate ubiquitous computing although we are not there yet – but we all have to prepare ourselves for it, both as customers and agents. And – naturally – it is about the boundaries of geometry and graphics or rather going beyond them.

2. THREE CESURAS IN THE EDUCATION OF AN ARCHITECT

Three fundamental genres of art – sculpture, drawing and painting – dominated the education of architects from the beginning of the 19th century. The first curricula were copied from academies of fine arts, augmented with descriptive geometry, and later completed with classes on construction material such as reinforced concrete.

The merger of architectural styles – the end of the 19th century became an eclectic hotspot –

with the evolution of building methods such as the elevator and the fireproof steel frame construction provoked programmatic and aesthetic manifestoes.

The Bauhaus marked a second fundamental cesura in the history of architectural education. Since the 1920's a triangle constituted modern architecture – the eternal quest for building materials to withstand gravity, the entitlement for aesthetic novelty derived from modern art and the constant push for innovation in the construction industry. The Weimar Bauhaus issued a new bilateral program based on the resentments of the Arts and Crafts movement: Thus the biased unity of conventional craftsmanship and progressive aesthetics such as *decor* less simplicity became the brackets for the education of architects.

A third cesura so far has not made it into broader education routines yet: The "European Agenda for Higher Education in the XXIst Century" (1997) calls for the transition "from teaching to learning" [1]. Normally, a teacher (singular) focuses on the basics of the trades *unilaterally* right along the line 'one has to learn the craft in order to proceed...' Now the question is: Who defines the basics in the information society where everything appears to be interconnected?

The transition "from teaching to learning" depicts nothing less than a paradigm shift within the Shannon and Weaver Model of Communication (1949). Now the receivers of the messages (plural) define their learning progress *individually*. That means, as Hans Ulrich Reck puts it, the "abandonment of polytechnic education in favour of experimental, exemplary and situative work". Reck urges to "accelerate the transitions in methodology" [2]. Such transitions will have to reach way beyond the established feedback models of communication.

3. CURRICULA FOR THE ARTISTIC FUNDAMENTALS IN THE AGE OF COMPUTERIZATION AND NEO-LIBERAL IDEOLOGY IN EDUCATION

I am in the process of developing a new curriculum for the artistic fundamentals in the architectural education at the Technische Universität Dresden. It is based on three main considerations:

(1) The need to teach more with less personnel due to substantial budget cuts: The percentage of students of an age group has increased by the factor of ten over the last 100 years in Germany. Now every second high school graduate moves on to university, due to the job market requirements for higher education. This increase triggers two oppositional developments. It spreads the students IQ, to use this rather politically incorrect formulation. Now it appears almost impossible to tailor classes for the first 50% of the population – their cognitive skills are just too dispersed [3]. Simultaneously the neo-liberal reconfiguration of education in Germany is taking off. Since the single biggest elevating screws for the ministries of education in the period of debt limits are the labor costs, they are gradually reducing the number of assistant professors and postponing or even preventing the replacement of professors who are eligible to retire. This way some departments disappear entirely, without much discussion about the importance and the substitution. How do we cope with such structural damage?

(2) The desire to balance and crossover analog and digital skills: While the computerization of the profession of the architect was quickly and thoroughly established in the 1990ies the education of an architect still labors on how to structurally and didactically implement Computer Aided Design, Building Information Modeling, Computational and Parametric Design, as well as graphic design and image processing [4]. Some schools of architecture offer profound curricula in Computational and Parametric

Design. Others revert to, or even outsource, software classes, sometimes directly to software firms with sales expectations. Quite often teachers and students alike assume that skills would somehow develop over night, without any support, by sheer coincidence. Some faculties are still subdivided in an analog fraction that eyes information society as a bad weather period, and a digital fraction that tries to get rid of such old-school stuff like free-hand drawing and Styrofoam models. Generally we endure the lack of profound surveys on the combination of analog and digital skills.

(3) The much needed adaptation of didactics to teach students how to learn and how to experiment. The education of an architect in the analog era appeared to be a tedious process. Even a terrific master thesis was far away from the benchmark of a professional working drawing. For most students a professional-looking presentation model was out of reach. Now, for almost a generation, even absolute beginners, freelancers and freshmen can produce professional-looking drawings, renderings and models. That became possible for two reasons: Digital media, from the start, allow for simple and sophisticated copy-and-paste procedures. Digital drawings, renderings and models look and are accurate, calculated by professional software. Teachers and students alike do get fooled by this. How can something not be right since it looks awfully accurate? Analog design work came by piles of tracing paper with an intrinsic mode of gradual development. Digital design work, at least for mediocre students, often appears as a successful start-finish affair with apparently little need for second-guessing, result control and alternative thought.

4. REAL LIFE AND REAL REPRESENTATION: EDUCATING ALONG THE BOUNDARIES OF GEOMETRY AND GRAPHICS

We are primarily fascinated with graphics and geometry in childhood. One of the most intriguing tasks for a child consists of the crayoning of contours, i. e. to apply graphic to

geometry, or pixels to vectors. Soon after, after what Betty Edwards calls “the stage of realism” by around age ten to eleven, the “crisis period” emerges: “a conflict between increasingly complex perceptions of the world around them and their current level of art skills” [5]. We don’t have to follow Edward’s adaptation of Roger W. Sperry’s research on human brain-hemisphere functions, the “L”- and the “R”-mode of drawing to accept that most people develop a clear preference for either a rather straight, mathematical correctness, i.e. geometry, or a rather “curvy” artistic ambiguity, i.e. graphics in general.

One of the fundamental didactic challenges persists in the combination of geometry and graphics since they often are performed in an either-or mode. There are the mathematically inclined students, and there are the artsy ones. We realize that the fundamental mental preference has been impressed in the curricula of schools (math and the arts) and universities ever since Immanuel Kant’s “The Contest of Faculties” (1798). The rift almost naturally extends into contemporary software branches where digital drawing as in Computer Aided Design actually means describing geometry by drawing lines, and digital image processing that has become synonymous with graphics although it is mostly photographs that are processed. Meanwhile talent has become a dubious, ambiguous term: It still refers to artistic talent and – that’s the novelty - the ability to work with computer programs.

5. THE FOUR-CORNER MATRIX

A contemporary didactic concept for the education of the architect along the boundaries of geometry and graphics has to be based on broad fundamentals.

The specs for these fundamentals sound universal: We want to deal with the real carbon world because that is where we live in and that is what we want to work for architecturally. We want to anticipate ubiquitous computing although we are not there yet – but we all have to prepare ourselves for it, both as customers and agents. And - naturally – it is about the

boundaries of geometry and graphics or rather exceeding the boundaries.

The didactic concept, therefore, consists of a matrix with four corners: Material world, ubiquitous computing, lines and color patches. Each category is described as a polarity. The material world consists of (a) real thoughts and (b) physical particles. Ubiquitous computing describes (c) specific action in (d) certain settings. Lines refer to (e) mathematical ideas and (f) general development. Color patches circumscribe the investigation of (g) images and (h) concepts (Figure 1).

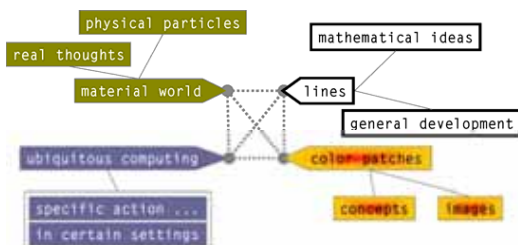


Figure 1: Four-Corner Matrix.

These are some of the challenges:

1. We have to cover wide arrays of topics both thoroughly – there are lots of specifics - and briefly at the same time, to maintain an overview with the sovereignty to navigate across the boundaries of the domain at hand. The nucleus of the new curriculum for the artistic fundamentals consist of a six-week course with approx. 150 students of architecture and landscape architecture in their third semester that combines various means of architectural and artistic representation like photography, painting, modeling, illustration, projection and graphic design. Each means is introduced as a magic square with nature, artistic intent, tools and didactics as the corners of the rectangle (Figure 2).

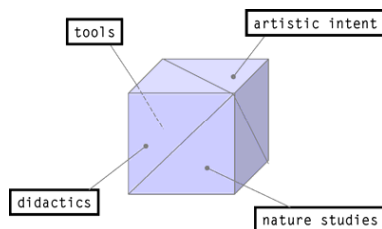


Figure 2: Magic cube made of nature, artistic intent, tools and didactics as the corners of the rectangle.

2. The didactics should aim at frequent intersections of the carbon world and virtual environments. Therefore, the nucleus of the new curriculum for the artistic fundamentals offers analogue and digital legs of the journey. It touches a variety of analogue techniques like painting, handicraft work and sketching, as well as digital image processing, desktop publishing and digital 3D modeling.
3. We have to focus on logical models and the ambiguities that come with boundaries between opposite genres. The new curriculum emphasizes that there are rules worth obeying, that we have to learn a craft in order to elevate it and that quite often there is no reasonable law to follow.
4. The didactics should foster the ping pong between two- and three-dimensional works. Eventually one source for mistakes is to treat architectural plans as 2D-images and not what they are - representations of spatial concepts. As opposed to that, it often proves helpful to “save” a 3D-model as a 2D-image in order to enhance its features. (“To save” can be understood in an analogue and a digital fashion). On top of everything we also enhance our cognitive capabilities by oscillating between images, concepts and spaces. It ought to improve our sovereignty in dealing with creative roadblocks and architectural bottlenecks, something that Russell Ackoff famously called “messes” [6].

5. Last but not least we want to encourage experiment while maintaining an encouraging set of rules and grades. Therefore, the course balances learning environments like excursions, field surveys, art room seminars, homework assignments and dedicated space for self-study.

6. THE CURRICULUM

The six-week course consists of three didactic layers: (1) The magic cube made of nature, artistic intent, tools and didactics as the corners of the rectangle (Fig. 2). (2) The combination of various means of architectural and artistic representation like photography, painting, modeling, illustration, projection and graphic design add practical depth to the matrix (Fig. 1). (3) In order to avoid jamming we offer two persistent columns of consideration. One column operates with an object of the real-life carbon-world (Fig. 3).



Figure 3: Indistinct, given objects (plural) in scale 1:1: Balance of traces of human effort plus political connotation: Learn that this object is a relict of destruction in the Second World War on display in the Dresden municipal lapidarium, the ruins of the Church of Zion, built between 1908 and 1912, destroyed during the allied firebombing 1945 and tentatively rehabed between 1994 and 1996.

The other column deals with a scaled sculpture model (singular) to depict architectural

connotations of local “high performance” architecture that the student created in one of our preceding classes (Table 1).

Table 1: Course outline.

Title	2D versus 3D	Given objects in scale 1:1	Scaled sculpture model
1. Photography + Digital image processing	3D → 2D	• • •	•
2. Analog painting	3D → 2D	• • •	• • •
3. Digital 3D-modeling	2D → 3D	•	•
4. Handicraft work	3D → 2D → 3D	•	•
Pictograms	3D → 2D	• • •	• • •
Projections	2D → 3D → 2D	- either or -	- either or -
5. Paper collage	3D → 2D	- either or -	- either or -
6. Digital poster	3D → 2D	- either or -	- either or -

The students are encouraged towards parallelisms of the sculptural objects from the environment and their sculpture model from the preceding class.

7. THE PRELIMINARY CURRICULUM

7.1 Photography + Digital Image Processing

Photographic snapshots come naturally to us. To succeed in documenting characteristic spatial features of an object under disadvantageous circumstances – insufficient illumination, partial overlaps with other objects, shadows, spider webs, dirt and chill – proves to be something else.

- Primary concern: How do we photograph objects to allow an optimum of spatial information? – Answer: With three sides visible.
- Secondary concern: How do we isolate the chosen object from its surroundings and to care for better brightness? – Answer: Use of sheets of paper (Figure 4a).
- Primary goal: Series of isolated photographic images that can be image processed.
- Secondary goal: Intuitive reinvention of two – three credible, convincing shades (natural light from different directions and angles).



Figure 4a-d – digital image processing. Ute Maria Hagelstein: In-situ photo work; series of digitally processed images, 3 of 9 (2013).

Common difficulties and frequent mistakes: Some students have problems to tap into their daily experience of how the sun produces shadows and are furthermore unable to apply their knowledge of shadow construction in descriptive geometry to the task at hand. Thus artifact and the shadows often do not appear authentic, consistent or appropriate (Figure 5a-e).

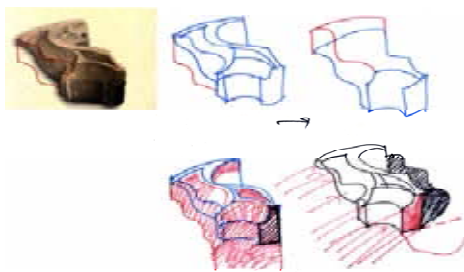


Figure 5a-e – variation of digitally processed images. Franziska Gernhard: Isolated relict; Niels-Christian Fritsche: Illustrations of alternative illumination scenarios and the reinforcement of contours (2013).

7.2 Analog Painting

Parting from the traditional artistic education of architects a few generations ago and the advent of the design with the computer sank analog painting into oblivion. Only recently traditional painting re-emerges in the arts, namely as the

New Leipzig School (*Neue Leipziger Schule*), and in architectural design, for instance in the shape of Stephen Holl's water colors. Oil and acrylic painting are, independent therefrom, great means for instant feedback with the potential of enormous individuality in correlation to an adequate material appeal of real world objects.

- Primary concern: Assure students that they can produce a series of three to five convincing paintings in a short period of time, approximately two hours (Figure 6a-c).
- Secondary concern: Encouragement to work from memory including simplification, alteration (no need for forensic exactitude) and reinvention.
- Primary goal: Introduction of analog painting with an intriguing haptic appeal as an effective, powerful and individualistic tool in art and architecture.
- Secondary goal: Comparison between digital image processing and analog painting.



Figure 6a-c – analog painting. Friederike Beye: Series of acrylic paintings on paper, 297 x 420 mm each, 3 of 5 (2013).

Common difficulties and frequent mistakes: Some students have problems to quickly prime their sheets, to fit the image onto the sheet, to explicitly distinguish three sides, and to apply a shadow to the image.

7.3 Digital 3D-Modelling from a Photo Image

The rationalization of modern architecture marginalized the design by sheer images as it was often practiced prior to the invention of the descriptive geometry and legally binding working drawings.

Digital image processing and rendering expertise raise the stakes in presenting architectural designs up to a point where designs are exclusively presented as naturalistic images without a thorough set of plans and sections. Such a design practice – based on a variety of software – encourages the reverse engineering of images into 3D models and subsequent sets of working drawings for building information modeling (Figure 7).

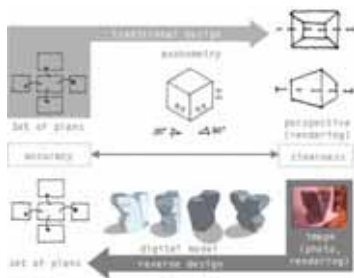


Figure 7 – designing back and forth. Traditional design starts from scaled 2D sets of plans in order to advance to a carefully constructed perspective drawing via axonometric; reverse design transforms an image into a sophisticated 3D model. Sketches: Grit Koalick (2004).

The task at hand encourages students to realize 2D images as active design sources rather than ultimately finished photos.

- Primary concern: Foster a straight visual, intuitive approach to transfer a 2D-image into a 3D-model via open source software like Google Sketch-up.
- Secondary concern: Establish the series motif – here: screenshots - to allow a maximum of spatial information with the

tendency to create the appearance of movement.

- Primary goal: Introduction of visual reverse engineering. How a design can be derived from an image as it is commonly done in contemporary architecture.
- Secondary goal: Exercise in spatial recognition of contours of irregular objects.

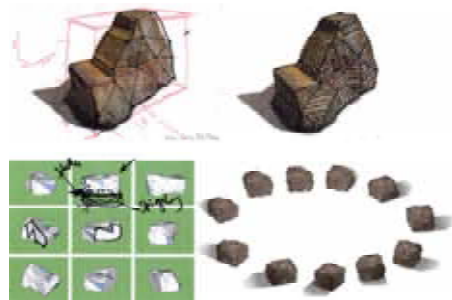


Figure 8a-d – digital modeling. Georg Hühn: Two sketches on ink jet prints; Lena Löhnert, Vivie Veith: Screen shots with 3D-modelling from a photo image, all 297 x 420 cm (2013).

Common difficulties and frequent mistakes: Some students have problems to distinguish three sides of their motif in order to come up with a base form.

7.4 Handicraft Work

One of the fundamental acceptance problems of contemporary architecture lies in the layman's inability to read architectural drawings as depiction of spatial information, i.e. concepts. One way to encourage spatial reading of 2D-images is to reconnect to handicraft work from childhood.

- Primary concern: Inspire spatial indication in 2D images in order to proceed from optional taste to mandatory information.
- Secondary concern: Differentiation between architectural working drawings addressed to experts and construction manuals for amateurs.

- Primary goal: Introduction of the correlation between a teaser 3D-image (this is what we want you to produce) and the 2D craftwork (that is how you start your work).
- Secondary goal: Layout training.

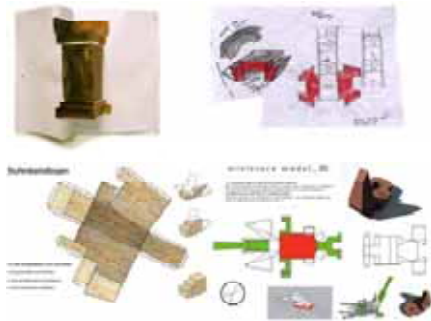


Figure 9a-d - Handicraft Work. Sophie Dietrich: “Faltburg” - folded acrylic painting, 297 x 420 mm; Max Oelke: Sketches on tracing paper, ~ 340 x 550 mm; Marlene Brugger, Anja Krüger: Papercraft sheets, 297 x 420 mm (2013).

Common difficulties and frequent mistakes: Some students have problems to switch from a 2D-image into spatial thinking. Some have problems to render their monolith buildable, i. e. to design the elevations of their paper model. Some fail to come up with a fitting title of their work or get lost in formulating the workflow.

7.5 Series of Deallocated Pictograms

Although and perhaps because human perception is based upon the recognition of simple signs, it turns out to be difficult to artistically operate with simple patterns – they already appear to be final and proof, therefore, difficult to operate with, without bias.

- Primary concern: Mêlée between lines and patches as well as two- und three-dimensional work motifs.
- Secondary concern: Fluency in switching between recognizable patterns – i. e. to design a logo for the object of origin - and deallocated forms.

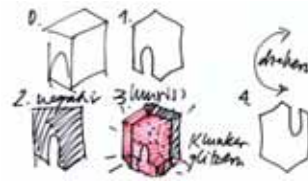


Figure 10 – designing and varying forms and signs into prospective pictograms: Outline, bisection, rotation, mirror, jewel, turn. Niels-Christian Fritsche: Introductory sketches (2011).

- Primary goal: Growing stock of abstract patterns.
- Secondary goal: Ingenious patterns for architectural design.



Figure 11a-d – Pictograms. Ute Maria Hagelstein: Signs. Felt pen on paper, 280 x 800 mm; Toni Emmrich: Pictograms. Felt pen on trace paper, 330 x 530 mm (2013).

Common difficulties and frequent mistakes: Students resort to a primary “bauhausesque” modus operandi with addition, rotation and mirroring instead of the design steps suggested: Outline, residual area, extrusion, jewel, selection and polarization.

7.6 Projections of the models on slanted surfaces (origami “A fortune teller”)

The preceding tutorials asked for series of works in the stringent mode of variations. Now, with the introduction of “projecting” a new intentional fuzziness is introduced: How to project one of the images from earlier exercises on the slanted surface of an arbitrarily chosen object, here an origami motif?



Figure 12 – Rows of projections. Seminar introduction: Handicraft work and projecting models on slanted surfaces. Niels-Christian Fritsche (2013).

- Primary concern: Training to project and to align 2D-images on 3D-surfaces
- Secondary concern: Accept that there is no accurate solution but approximation on the basis of trial and error.
- Primary goal: Accept that there is no accurate way to come up with the ultimate solution. This would require a technical setup with a slide projector or a beamer and a camera...
- Secondary goal: Alternative branch of abstract patterns.

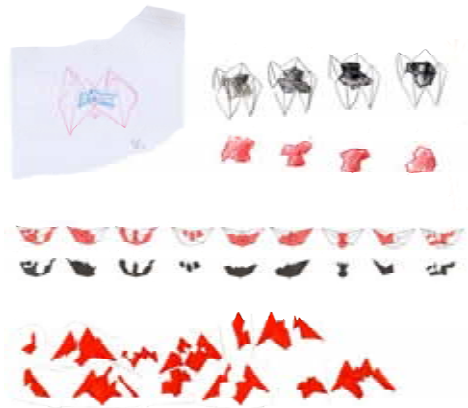


Figure 13a-d - Projections of the models on slanted surfaces. Rosa Lynn Pakusch: Pen on trace paper, 330 x 340 mm; Nina Schaar: Colored pencil on paper, 297 x 4230 mm; Michael Penseler: Felt pen on paper, 120 x 870 mm; Farhad Babayev: Felt pen on paper, 180 x 710 mm (2013).

Common difficulties and frequent mistakes: Some students confuse projection with the overlay of two images. Some students reject the fuzziness of the task altogether. Some are unable to isolate the projection from the surface they are projected on.

7.7 Paper collage

The large-scale format of the poster reapplies the projection method by reversing 3D images into a collage array of interdependent projections as autonomous visual areas.

- Primary concern: Alignment of the collage of one of the given objects with projections of the objects in four directions.
- Secondary concern: 100% coverage with a balance between then signature objects and the background.
- Primary goal: Projection of casual collage material onto a well-known motif.
- Secondary goal: Realization of the collage as a fundamental design tool.

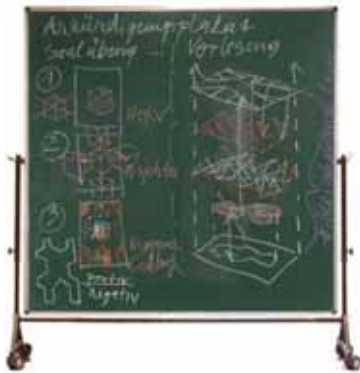


Figure 14 – Seminar introduction. Left column: Paper collage in three steps: Apply random collage material to motif, project motif to all sides including bottom view, and fill in of the in-between spaces. Right column: Layering as it is applied in image processing software. Niels-Christian Fritsche (2013).



Figure 15a-b – Paper collage. Frances Ebert: Collage (colored paper, Japan paper, rice grains, newspaper, duct tape), 1000 x 700 mm; Shiva Franizari: Collage (mixed media), 1000 x 700 mm (2013).

Common difficulties and frequent mistakes: Some students have ongoing problems in rotating the motifs they have been working with for a number of weeks on a given sheet of paper. They often try to imitate the appeal of the original motif with random collage material instead of speeding off. Some students do not

let go of the object reference in favor of a new graphic product.

7.8 Digital Poster

The evolution of poster design indicates the limitations of layers, both traditionally and in the digital age: Text is usually applied on top of an array of images.

Here, the aim is to differentiate between the strengths of the *intuitive* paper collage and *structured* digital layering. How can the richness of the paper collage be transferred but not copied into the digital realm? Again, no assignment topic is given; students are encouraged to create their posters from scratch and to design without bias.



Figure 16a-k – components of the poster. Gauges, fonts [7], hand-made and digital graphic structures (hand made paper, whisked QR-code).

- Primary concern: Introduction of contemporary graphic design as the parallelism of photorealistic images, stencils and signature fonts.
- Secondary concern: Weaving of the components in opposition to the common vertical layering in graphic design.
- Primary goal: Acceptance of the visual race between three different graphic branches and the visual limitations of software products.
- Secondary goal: Acceptance of the balance between visual perceptibility and graphic appeal.

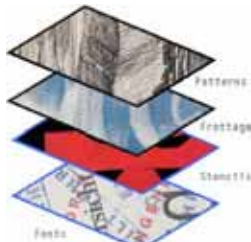


Figure 17 – Layer structure of the ubiquitous image processing software, to be weaved [8].

- Primary goal: Demonstrate and overcome the limitations of image processing and layout programs that are organized vertically along separate layers by the introduction of weaving modes.
- Secondary goal: Illustrious graphic product with “whoa!”-factor.



Figure 18a-b – Digital poster: Jonas Buddemeier, 600 x 840 mm (det., 2012); Isabella Dahm, 320 x 450 mm (det., 2013).

Common difficulties and frequent mistakes: Some students confuse the digitalization with a simple scan copy of their collage. Often the product still displays traditional vertical layering with little or no weaving of the components.

8. CONCLUSIONS

In my paper I suggest exploring exemplary didactic approaches in analogue and digital fashion, as well as 2D versus 3D, a ping pong of geometry, pictures, images and models, with the chance to experimentally cross over from one realm into the other.

The physical goal is a portfolio of what I call “direct graphic design”, i. e. graphic design that is produced immediately with the work in front of the producer, where action starts from scratch, and where the proceeds are constantly echoed by immediate response by the proceeding work of the artist, and with little or no guidance by classic rules of the trade like theory of colors, the golden section and typography (they all appear in different classes).

The cross-over of categories that normally are thought to be separate is strongly encouraged for a number of reasons:

First of all speed. The course shall inspire students to work in series to allow comparison and approximation: Which one works best?

Second: The variety of techniques and software branches. We all have our either-or preferences. Productive mistakes and innovation often occur on the edges and in between fields. In order to get to the edges we almost always have to focus on at least two techniques and / or fields of expertise to encourage experiment.

Third: The cognitive training that comes with the mêlée between two- and three-dimensional works that characterizes the work of the architect.

Fourth: The start of a personal library of universal graphic patterns that are not tied to specific tasks. Instead of pure problem solving the start of an individual all-across body of work is suggested.

Fifth: Rules are a powerful help but no stand-alone guarantee for success. In order to proceed we have to learn the trades, and then to set them aside.

Sixth: Studying architecture all too often appears as a sequence of unrelated fields of study. We want to bring the opposites together in order to instigate integrative thinking.

ACKNOWLEDGMENTS

The author is deeply grateful for the seminal support of the assistant professors Dr.-Ing. Tom Schoper, Dipl.-Ing. Till Schuster and Dipl.-Ing. Felix Greiner Petter in the realization of the

large-scale seminars at TU Dresden. Gabriele Thierfelder has been pivotal in organizing and documenting the stream of thousands of images.

The permission given by the Amt für Kultur und Denkmalschutz, Gerd Pfitzner and Dirk Schumann, to visit the, and to photograph in the Dresden municipal lapidarium, the ruins of the Church of Zion was fundamental in establishing the course.

REFERENCES

- [1] European University Association (EUA) / CRE and UNESCO-CEPES 1997: "European Agenda for Higher Education in the XXIst Century", Palermo, in preparation of the UNESCO-World Conference 1998.
- [2] H. U. Reck. Zwischen Bild und Medium. Zur Ausbildung der Künstler in der Epoche der Techno-Ästhetik. In: P. Weibel (Ed.): Vom Tafelbild zum globalen Datenraum. Neue Möglichkeiten der Bildproduktion und bildgebender Verfahren. Hatje Cantz, Ostfildern-Ruit 2001, p. 17-49, p. 28. Translated by the author.
- [3] E. Stern, A. Neubauer. Nature via Nurture. Warum eine Universität für alle niemandem nützt: Intelligenzunterschiede lassen sich nicht reduzieren. In: *Forschung & Lehre* 8-2013, pp. 634-636.
- [4] D. Lordick. Call for papers: *Beyond rendering berlin 2012*, 8th Conference of the Deutsche Gesellschaft für Geometrie und Grafik (DGfGG), Berlin, May 16-18, 2012: <http://www.brb2012.de/?q=Tagungsthema>.
- [5] B. Edwards. The New Drawing on the Right Side of the Brain. Putnam, New York 1999, p. 69-81.
- [6] R. Ackoff. The Future of Operational Research is Past. *Journal of Operational research Society*, 30 2 (Pergamon Press, Ltd., 1979), p. 93-104. Quoted after D. A. Schön: The Reflective Practitioner: How Professionals Think in Action. Basic Books, New York 1984, p. 16.
- [7] Fonts: <http://www.1001freefonts.com> (2011-2013). 96 Surface Design Group/SDG: Acidic (1995), Typeface: Broken Fluid (2011), <http://truefonts.blogspot.com> (kaiserzharkhan): Escanded (2012), Michael Tension: Impact Label (2008), junkohanhero: Phorssa (2007), Måns Grebäck: Gretoon Highlight (2011).
- [8] Layer motifs: S. Beyder, B. Morlock: Kreativer Kunstunterricht in der Sekundarstufe. Auer, Donauwörth 1996, p. 23, 131, 133 (modified). Top image after Max Ernst: Historie naturelle. Photoengraving after pencil frottages, 51,7 x 35cm, map, sheet. XVIII: Les mœurs des feuilles („Blättersitten“).

ABOUT THE AUTHOR

1. Niels-Christian Fritsche is professor for architectural delineation and presentation methods at the Technische Universität Dresden.

A RECYCLABLE MÖBIUS STRIP

Martin PFURNER, Josef SCHADLBAUER and Hans-Peter SCHRÖCKER
University of Innsbruck, Austria

ABSTRACT: This paper deals with one of the most famous mathematical objects, the so called Möbius strip. The original example proposed by A. F. Möbius is made from a slip of paper and is thus developable. Hence, there has always been a certain interest in developable examples. We give a historical overview on the contributions to this topic. It has been said that the famous recycling symbol is derived from a Möbius strip. This is, indeed true for some variants of the symbol. In our contribution we connect the recycling symbol to a special Möbius band, composed of pieces of circular cylinders and planes and exhibiting a three-fold symmetry. Its orthographic projection in the direction of the symmetry axis gives the recycling symbol. We provide a detailed construction for this object and describe its didactical values.

Keywords: developable Möbius strip, recycling symbol, 3D-CAD, descriptive geometry, self-supporting structure

1. INTRODUCTION

The aim of this paper is twofold. We want to give an overview of the fascinating history of developable Möbius strips, starting with the days of F. A. Möbius and ending with nowadays numerical solutions to the variational problem that determines the stable shape of the strip (Section 2). This history is full of remarkable contributions by individuals, often well ahead of their times, and ultimately lead to a satisfactory theoretical solution. Unfortunately, the numeric solution for the stable shape of the developable Möbius strip and other similar shapes is still only accessible to experts while there is certainly a wider need in architecture and design communities.

In Section 4, we continue by investigating the relation of a particular developable Möbius strip and one of the most well-known logos, the universally used recycling symbol. Our interest was actually stirred by a wrong interpretation of a remark by W. Wunderlich in [?]. Wunderlich describes an example given by M. Sadowsky in [4] – a paper that was not accessible to us at that time. As our suggested model, this developable Möbius strip consists of parts of planes and right

circular cylinders. But, in contrast to our example, it lacks rotational symmetry and equal radii of cylinders. As a spatial structure, it is rather trivial. Sadowsky presented it to silence doubts on the existence of developable Möbius strips and to give an upper bound for the bending energy of the strip's stable shape.

Some variants of the recycling symbol (but not the original design!) can be seen as the orthographic images of a Möbius strip with three fold rotational symmetry. We suggest an elementary spatial construction for this Möbius strip, based on a self-supporting structure of three mutually tangent cylinders of revolution. This constitutes a nice modeling exercise that requires spatial abilities, spatial reasoning, and a good command of 3D-CAD software. The obtained solution can be the basis for an animated recycling symbol with arrows flowing around the Möbius strip.

2. HISTORY

In 1858, the German mathematician and astronomer F. A. Möbius presented his famous model of a one-sided surface, later denoted as Möbius strip. It can be made from a long rect-

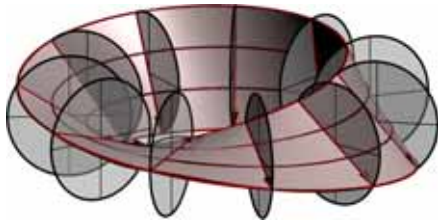


Figure 1: A rotoid-helicoid as topological Möbius strip.

angular slip of paper by gluing together opposite vertices after a twist through 180° about the longer mid-line. The popularity of the Möbius strip is due to its curious topological properties – it is a one-sided surface with a boundary that is homeomorphic to a circle – but certainly also due to its simple fabrication.

It is easy to find parametric surfaces that are topologically equivalent to Möbius' surface. One example, a rotoid helicoid [1, Ch. 7], is depicted in Figure 1. We may think of this surface as being generated by a diameter of a circle that rotates at constant angular velocity ω about its axis while at the same time undergoing a rotation of angular velocity 2ω about a second, co-planar axis.

In contrast to Möbius' example, the surface in Figure 1 is not developable. This can be inferred from the slightly curved silhouette. Until 1930, when M. Sadowsky [3] provided the first and elementary example, the theoretical existence of developable surfaces with the topology of the Möbius strip was unclear. But Sadowsky's most important contribution in [3] was the description of the surface energy (the integral over the square of the non-vanishing principal curvature) that is minimized by Möbius' original paper example. If fabricated from sufficiently stiff material, the paper slip will attain a stable shape that is a local minimizer of this energy. Of course, at that times, a solution to this variational problem was completely out of reach.

This motivated W. Wunderlich in 1962 to describe an approximate model [8]. Wunderlich



Figure 2: Wunderlich's developable Möbius strip

carefully examined the symmetries of a stiff paper model and, adopting Sadowsky's simplifying assumption of infinitesimal strip width, came up with a suggestion for the strip's mid-line. The Möbius strip is then part of the mid-line's rectifying developable (the hull surface of the planes spanned by the midline's tangents and bi-normals). Wunderlich's ingenious modeling of the midline as rational sextic without modern spline and free-form techniques is quite remarkable and computer graphics visualizations of his example look convincing (Figure 2).

In more recent papers, [2] and [6], the equilibrium equations for Möbius' strip are solved numerically. Especially the later paper received wide appreciation and was discussed in quite a few popular media. It turned out that Wunderlich's example comes close to reality for small band width but fails for broad bands where creases occur. The numeric solutions probably mark the end of efforts to find the true shape of Möbius' surface. What is still missing is a robust and user-friendly integration in popular CAD systems. Equilibrium shapes of flexible material under some additional constraints are actually frequently asked for by architects and designers who want to transfer their prototypes into CAD models. Here, the discrepancy between the sim-

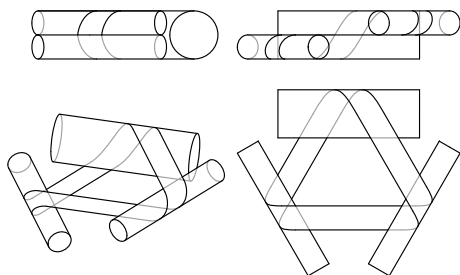


Figure 3: Sadowsky's example of a developable Möbius strip

ple fabrication and difficult modeling becomes apparent.

An account of the developments before 1990 is given in [5]. There the author also re-invented Sadowsky's first developable example from [3] which is briefly described by Wunderlich in [8]:

- It consists of three cylindrical and three planar parts and
- the three cylinder axes enclose angles of 60° .

The rather trivial example of Sadowsky (and later Schwarz) is depicted in Figure 3. The three cylinders have respective radii r , r , and $2r$; their axes are parallel to a common plane as are the three planar parts.

3. A DEVELOPABLE MÖBIUS STRIP AND THE RECYCLING SYMBOL

This is, however, not the only possible interpretation of Wunderlich's description. Not having access to Sadowsky's original paper (now we do) we thought of a structure of three cylinders with mutually *skew* axes and a three fold rotational symmetry. This interpretation is supported by the ubiquitous presence of symbols or logos based on a Möbius strip wrapped around three cylindrical shapes, for example the logo of the German Mathematical Union (Figure 4) or part of the logo for the Möbius ProjectTM from Maplesoft (Figure 5).

The most prominent example among these is logos and symbols is certainly the recycling symbol (Figure 6, top). The original design from

1970 is due to G. Anderson. This logo itself is in the public domain. Therefore, numerous variants exist and are universally used and understood. Different versions of the recycling symbol even have entries in the Unicode standard (Figure 6, bottom).

Note the lacking rotational symmetry in Anderson's original design when compared to the Unicode symbols. The original logo consists of three arrow parts, one leading from top to bottom and two leading from bottom to top, very much like Sadowsky's and Schwarz's model but with three equal radii. Thus, it is difficult to view it as image of a spatial structure.

It is our aim in this paper to describe a developable Möbius strip with three fold rotational symmetry which is composed of three cylindrical and three planar pieces in such a way that its orthographic projection into the direction of the axis of symmetry can serve as a basis for a recycling symbol. One motivation for this is the possibility to animate the three arrows of the recycling symbol as they move along the band. Once the 3D-model is available, this is fairly straightforward. The actual construction is a demanding task in spatial modeling but might be interesting from a geometrical and didactic viewpoint. In this sense, it can be seen as an extension of the self-supporting structure that was presented and discussed at the ICGG conference in 2008 [7, 9]. Incidentally, our construction is based on precisely this self-supporting structure.

4. CONSTRUCTION OF THE MÖBIUS STRIP

For the construction of such a Möbius strip with a three fold rotational symmetry we use a self supporting structure with three congruent cylinders of revolution. Each cylinder touches the other two and the structure has such a three fold rotational symmetry as shown in top, front, right and an axonometric view in Figure 7. Mutual tangency of the cylinders is *necessary* as piece-wise planar parts in the Möbius strip must belong to a common tangent plane of two cylinders. Con-



Figure 4: Logo of the Deutsche Mathematiker-Vereinigung (used by permission)



Figure 5: Part of the logo of the Möbius Project™ from Maplesoft (used by permission)

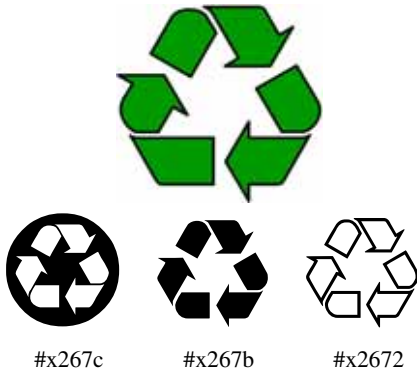


Figure 6: Anderson's universal recycling symbol (top) and recycling symbols in Unicode (bottom)

structions for this structure have been described in detail in [7]. We note that the construction is rather easy when the cylinders' radius is not given initially but is determined as half of the perpendicular distance of the three cylinder axes.

We start with some initial observations. The touching points of the three cylinders of revolution form a regular triangle t . Because of the symmetry of the self supporting structure, an altitude of this triangle and the two generators of the touching cylinders through the same touching point P lie in one plane. Therefore it is possible to find a strip in the common tangent plane in P of the two cylinders that lies symmetrically to the altitude of t through P , that is, the edges of the

strip lie perpendicular to the altitudes of t . This is necessary for the rotational symmetry of the Möbius strip.

The part of the strip that bends around the cylinders of revolution should be a rectangular strip in the unrolling. Therefore, it is bounded by two helical curves on the cylinder.

It is easy to complete the construction of the complete strip (and the derived 3D recycling symbol) if we can construct its spine curve (mid-line). Therefore, we only describe the construction of the mid-line. It consists of straight lines touching two adjacent cylinders. These line segments are connected by helical curves on the cylinders with tangent continuity. In the unrolling of the strip, the mid-line maps to a straight line.

We can start with an arbitrary straight line from one point of the altitude of t orthogonal to it and tangential to one of the cylinders that touch in the corresponding corner of t (see Figure 8). Note that this line only gives the direction of the spine curve. It will be translated later in order to ensure rotational symmetry.

To construct the helical curve on the cylinder of revolution with this start tangent direction we use the angle preserving property of unrolling. This mapping preserves the angle between the start tangent and the generating lines of the cylinder. In the unrolling of the cylinder the helical curve becomes a straight line with the same direction as the start tangent line. If we use the

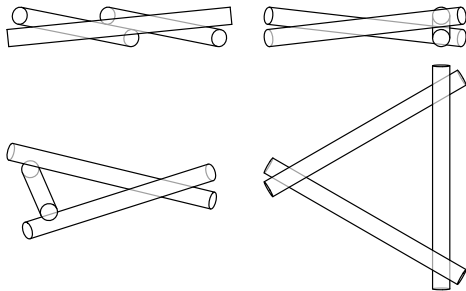


Figure 7: Self supporting structure

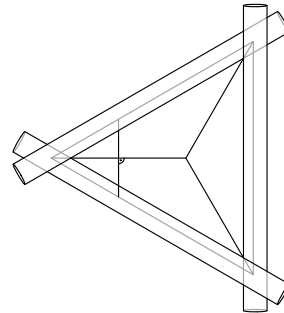


Figure 8: First construction step in top view

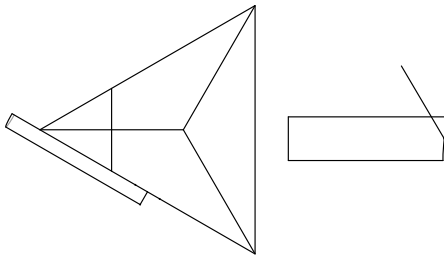


Figure 9: Construction of helical curve in un-rolling

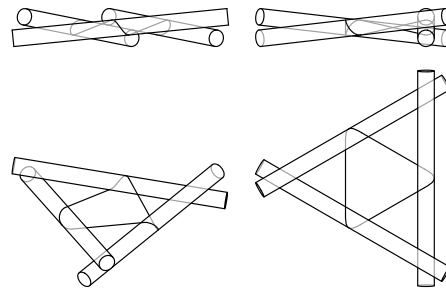


Figure 10: Spine curve of the Möbius strip in top, front, side and orthographic view

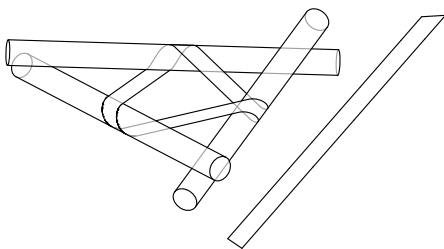


Figure 11: Möbius strip with three fold symmetry and its unrolling

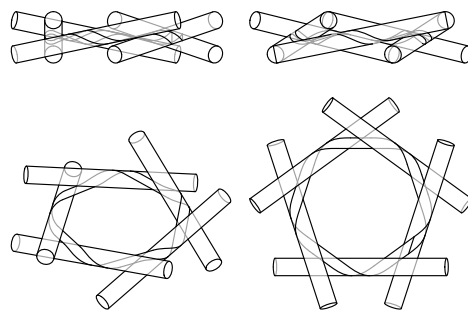


Figure 12: Möbius strip with 5-fold rotational symmetry

unrolling of the part of the cylinder between the two generators of the touching points we can directly construct the height of the helical curve segments in direction of the cylinder's generating lines (Figure 9).

With this length it is possible to construct the helical curve on the cylinder in such a way that it is symmetric with respect to one altitude of triangle t in the top view. The end points of the helix give the start points of the straight line segments and the rotational symmetry gives the whole spine curve of the Möbius strip (Figure 10).

The Möbius strip itself consists of planar parts that lie in the common tangent planes of the cylinders of revolution which are bounded by straight lines parallel to the spine curve. The part of the strip bending around the cylinders are bounded by pieces of helical curves to be congruent to the helical parts of the spine curve, as shown in Figure 11.

One possible recycling symbol based on the above mentioned construction can be seen in Figure 13. It has the three fold rotational symmetry that is lacking in Anderson's original.

Note that the construction can also be applied to self supporting structures consisting of $2n + 1$ cylinders ($n \in \mathbb{N}$). This leads to a Möbius-like strip with $2n + 1$ fold rotational symmetry. An example with a 5-fold rotational symmetry is shown in Figure 12. A corresponding recycling symbol with this rotational symmetry is shown in Figure 14.

5. CONCLUDING REMARKS

We gave a short overview of the history of developable Möbius strips and discussed in detail a discussion of a developable Möbius strip with three fold rotational symmetry that consists of pieces of planes and cylinders of revolution. This object can serve as the basis for constructing a 3D version of the famous recycling symbol.

We believe that the presented example exhibits many virtues that make it suitable for advanced CAD courses. It is related to familiar and universally known shapes, the Möbius strip and the

recycling logo, and it promotes a deeper understanding of both.

The outlined construction is certainly advanced but can be broken down to manageable steps with inherent didactic value. It touches topics like rotational symmetry and its exploitation for construction, the distance of skew lines in space and the tangency of cylinders. Moreover, we use two important techniques of spatial problem solving, the reduction of dimension by unrolling and the reduction to direction and angles (restriction to a spherical image) – a very common technique in spatial geometry and kinematics. All of this can be explained and trained at hand of the proposed example.

REFERENCES

- [1] G. Glaeser and K. Polthier. *A Mathematical Picture Book*. Springer, 2014.
- [2] L. Mahadevan and J. B. Keller. The shape of a mobius band. *Proc. R. Soc. Lond. A*, 440(1908): 149–162, 1993.
- [3] M. Sadowsky. Ein elementarer Beweis für die Existenz eines abwickelbares Möbiusschen Bandes und Zurückführung des geometrischen Problems auf ein Variationssproblem. *Sitzungsber. Preuss. Akad. Wiss.*, 22: 412–415, 1930.
- [4] M. Sadowsky. Théorie der elastisch biegsamen undeformbaren Bänder mit Anwendungen auf das Möbius'sche Band. In *Proc. Int. Cong. appl. Mech.*, volume 2, pages 444–451. Stockholm, 1930.
- [5] G. Schwarz. The dark side of the Moebius strip. *Am. Math. Monthly*, 97(10): 890–897, December 1990.
- [6] E. Starostin and G. V. der Heijden. The shape of a Möbius strip. *nature materials*, 6, August 2007.
- [7] K. Suzuki and H.-P. Schröcker. Application of descriptive geometry procedures in solv-

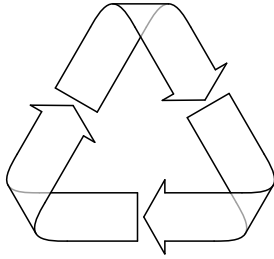


Figure 13: A recycling symbol derived from our construction

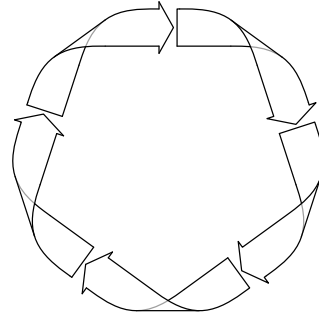


Figure 14: Recycling symbol with 5-fold rotational symmetry

ing spatial problems with feature and parametric modelling 3D-CAD. In *Proceedings of the 13th International Conference on Geometry and Graphics (ICGG 2008)*. Dresden, Germany, 2008.

- [8] W. Wunderlich. Über ein abwickelbares Möbiusband. *Monat. Math.*, 66: 276–289, 1962.
- [9] P. Zsombor-Murray. Tangency among three cylinders, a hyperboloid and a torus. In *Proceedings of the 13th International Conference on Geometry and Graphics (ICGG 2008)*. Dresden, Germany, 2008.

ABOUT THE AUTHORS

1. Martin Pfurner is Assistant Professor at the Unit for Geometry and CAD, Faculty of Engineering Science, University of Innsbruck, Austria. His research interests include classical geometry, didactics of geometry, kinematics and robotics. He can be reached by email via the address martin.pfurner@uibk.ac.at or through the postal address University Innsbruck, Unit for Geometry and CAD, Technikerstraße 13, 6020 Innsbruck, Austria. His homepage is <http://geometrie.uibk.ac.at/pfurner/>.

2. Josef Schadlbauer is Assistant Professor at the Unit for Geometry and CAD, Faculty of Engineering Science, University of Innsbruck, Austria. His research interests include classical geometry, kinematics and robotics, and line geometry. He can be reached by email via the address josef.schadlbauer@uibk.ac.at or through the postal address University Innsbruck, Unit for Geometry and CAD, Technikerstraße 13, 6020 Innsbruck, Austria. His homepage is <http://geometrie.uibk.ac.at/schadlbauer/>.

3. Hans-Peter Schröcker is Associate Professor at the Unit for Geometry and CAD, Faculty of Engineering Science, University of Innsbruck, Austria. His research interests include classical geometry and applications, applications of algebra to kinematics and robotics, didactics of geometry, and convex and discrete geometry. He can be reached by email via the address hans-peter.schroecker@uibk.ac.at or through the postal address University Innsbruck, Unit for Geometry and CAD, Technikerstraße 13, 6020 Innsbruck, Austria. His homepage is <http://geometrie.uibk.ac.at/schroecker/>.

THE REPRESENTATION ON THE CYLINDRICAL ROTARY SURFACE – INVERSE PANORAMA

Jolanta DŹWIERZYŃSKA

Rzeszow University of Technology, Poland

ABSTRACT: The subject of this paper refers to the author's previous study dealing with the possible application of computer software for direct and practical mapping of cylindrical and conical panoramas, as well as panoramas which backgrounds are approximated by right polyhedra. Spreading this idea, the author takes into consideration the projection onto the cylindrical, rotary surface (or on the fragment of this surface), in which the center of the projection is not located like in a typical panorama "inside" the surface, but outside of it. This representation is defined as a multicenter projection from the centers dispersed on a circle. Such approach gives maximum approximation of the given results of the projection to real perception one experiences observing the image. The graphical mapping of the effects of the representation is realized directly on the unrolled background of the projection. That is due to the projective and graphical connection between points displayed on the cylindrical background and their counterparts received on the unrolled, flat background. The paper shows descriptive way of creating such images. However, for the significant improvement of construction of lines, analytical algorithms in Mathcad program are formulated. They can be implemented in majority of the computer graphical packages, which makes drawing more efficient and easier.

The presented representation, and the way of its mapping on the flat background can find application in advertisement and art, when drawings are displayed on the cylindrically curved surfaces.

Keywords: Cylindrical Perspective, Panoramic Projection, Computer Aided Design.

1. INTRODUCTION

The paper is the continuation of the author's previous considerations dealing with panoramic projections onto cylindrical and conical rotary surfaces, as well as the direct way of mapping of them on the unrolled cylindrical background [4]. The theory of the panoramic presentation was developed in two directions distinguishing a single center panorama and a multicenter panorama [3]. This idea is expanded in the paper as it discusses the central projection from dispersed centers (multicenter panorama) onto a cylindrical rotary surface. However, the centers of the projection, contrary to the previous study, are located not inside the cylindrical surface but outside of it. Such panorama is called a inverse

panorama.

The use of the moving center of the representation aims at better approximation received projections of the given figures to the real images experienced while watching panorama.

One can find some examples of reverse/inverse panoramas in Internet [5],[6]. The greater and greater interest in such presentations is mainly initiated by the development of digital panoramic techniques: simple digital camera stitching methods and high-level panorama scanning [1],[2]. The analysis of geometrical aspects of these kind of presentations is a purpose of the paper.

2. BASIC INFORMATION

Like in case of linear perspective on a projection plane, considered representation is a two projective partly compound representation. Therefore, the mapping of any figure is composed of two projections of it: the *main* one and the *auxiliary* one. Moreover, it has the property, that the given figure and additionally the image of this figure got earlier in the auxiliary projection are represented in the main projection [3],[4].

Thus, the apparatus of the inverse panoramic projection is composed of:

- the apparatus of the main projection, which is the central projection from the center S_X located on the circle \hat{s} onto a rotary cylindrical background $\hat{\mathcal{L}}$
- the apparatus of the supplemental projection, which is the projection from the center $O \in \hat{\mathcal{L}} \perp l$ onto the projection plane π , which is perpendicular to the axis l (Fig.1a,1b)

The circle \hat{s} , which plane is parallel to the projection plane π , is regarded as the circle of the viewpoints.

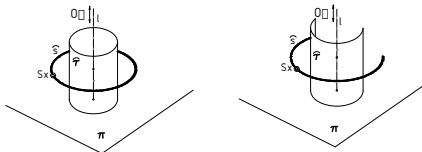


Figure 1: The structure of the apparatus of the representation.

The changeable center (viewpoint) S_F of the inverse panoramic projection is attributed to the given point F by cutting the circle \hat{s} by a half-plane determined by the axis l and the point F (Fig.2).

Due to this fact, the image of any point F , in general, is a pair (F^S, F^{OS}) of two projections included in the ruling t_F of the cylindrical surface. The point F^S is a main central projection from the centre S_F onto a cylindrical background $\hat{\mathcal{L}}$. Whereas, the point F^{OS} is a auxil-

iary projection. It is received by submitting the point F^O (the orthogonal projection of the point F onto the projection plane π) to the main projection from the centre S_F onto cylindrical background $\hat{\mathcal{L}}$ (Fig.2).

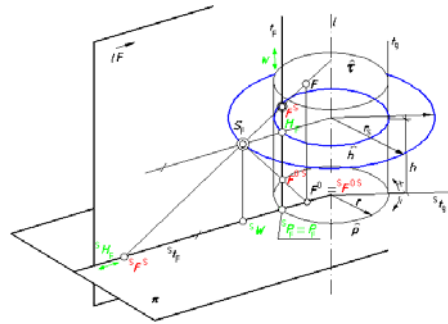


Figure 2: The projection of the straight line t_F and the series $t_F(P_F, H_F, W_F, \dots)$ of points for the realization of the transformation.

3. MAPPING OF THE INVERSE PANORAMA DIRECTLY ON UNROLLING BACKGROUND

For the graphical mapping of effects of the representation on the flat unrolled surface it is necessary to transform the images contained in the background $\hat{\mathcal{L}}$ to their counterparts on the unrolled background $\hat{\mathcal{L}}^R$. It is realized by the projection of each ruling t_F of any point F from the center S_F onto the projection plane (base plane) $\hat{\mathcal{L}}$, as well as establishing projective relations between points of the ruling S_{t_F} after that projection, and their counterparts on unrolled background $\hat{\mathcal{L}}^R$. Similar approach was presented in [4].

3.1 Mapping of a point

Geometrical action which permits to find the inverse panoramic projections on the unrolled surface looks in following way. First, the apparatus of the representation should be defined within an accuracy of isometry. Next so-called a base ruling t_g and the ruling t_F which contains a pair of projections of the certain point is dis-

tinguished in the cylindrical background(Fig.2). Moreover, the positive circulation for measuring angles \angle of the rotation transforming the ruling t_g into the ruling t_i is established. Next, the fragment of the unrolled background \square^R with distinguished: the base ruling t_g^R and the ruling t_F^R , as well as the base circle \hat{p}^R and the horizon circle \hat{h}^R are drawn in the projection plane \square . The image \square^R of the unrolled background \square is placed itself towards the base circle \hat{p} the way shown in the fig.3.

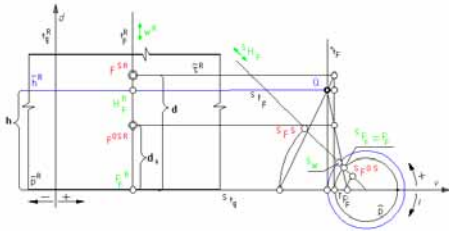


Figure 3: The way of the assignation of the homologous points of the series $t_F^R(W^R, P^R, H^R \dots)$ to the series ${}^S t_F({}^S W_F, {}^S P_F, {}^S H_F \dots)$.

The series $t_F(P_F, H_F, W_\infty, F^S, F^{OS}, \dots)$ of points belonging to the ruling t_F is distinguished; where the point P_F is included in the base circle \hat{p} , the point H_F is included in the horizon circle \hat{h} and the point W_∞ is a vertex of the surface (Fig.2). Next, the ruling t_F with the series of points established is projected from the centre S_F onto the projection plane \square . As a result of that projection the series ${}^S t_F({}^S P_F, {}^S H_F, {}^S W, {}^S F^S, {}^S F^{OS}, \dots)$ of points is obtained in the ruling ${}^S t_F$. Then, the ruling ${}^S t_F$ with the series ${}^S t_F({}^S P_F, {}^S H_F, {}^S W, {}^S F^S, {}^S F^{OS}, \dots)$ of the points included in it is turned around the centre of the base circle to the position it overlaps with the unrolled base circle \hat{p}^R (Fig.3). In a row, the ruling t_F^R matching the t_F one with distinguished series $t_F^R(P_F^R, H_F^R, W_\infty^R, \dots)$ of points is translated to the position of the straight line ${}^* t_F$. That geometrical action points out that series of points obtained after the rotation and included in the ruling ${}^S t_g$ and series of points

obtained as a result of translation and included in the straight line ${}^* t_F$ are projective. As they have also united homologous points they are perspective ones. It gives graphical relationship between the image of the point get in the additional projection from the centre S_F onto the projection plane \square with its counterpart get in the unrolled background \square^R and on the contrary. It is worth noticing that for the point F included in the projection plane \square is as follows: $F^S = F^{OS}$ and $F = {}^S F^S = {}^S F^{OS}$.

3.2 Mapping of a straight line

The main inverse panoramic projection of the vertical straight line in space is a vertical ruling of the cylindrical surface. Whereas, the main inverse panoramic projection of the horizontal straight line contained in the plane of circle \hat{s} is a horizontal circle \hat{h} . In case of the straight line not particularly situated towards the apparatus of the representation both projections: the main one and the auxiliary one are curved lines. They are the common part of the conoid and the cylindrical surface so not easy to construct in a descriptive way. The figure number 4 shows an example of the inverse panoramic projection of the square $\langle ABCD \rangle$ contained in the projection plane \square .

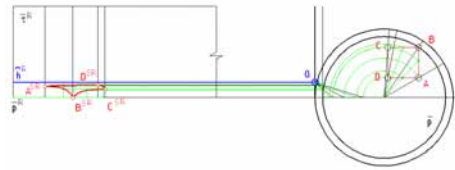


Figure 4: The inverse panoramic projection of the square $\langle ABCD \rangle$ contained in the projection plane \square .

In general the image of any figure in inverse cylindrical panoramic projection depends on the structure of the apparatus of the projection, as well as the location of the figure towards the apparatus.

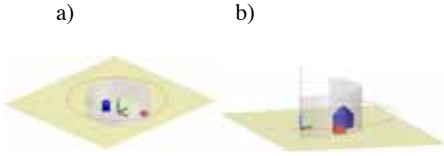


Figure 5: The possible structure of the apparatus of the inverse panorama: a-full panorama, b-part panorama

When the background is a full surface and the figure represented has the common points with the axis l the full panoramic image of the figure can be achieve (object panorama) (Fig.6).

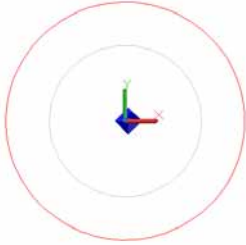


Figure 6: The location of the figure towards the apparatus for receiving an object panorama.

The object panorama cannot be achieved in case of the panoramic projection onto a part of the cylindrical surface. According to the projection definition only the points which distance from the surface is less than the radius of it can be represented. It is possible to determine similar representation from dispersed centers for the points which are situated further from the surface than the radius of it (Fig.7a). In that case, however, a reverse image of the represented figures can be achieved (Fig.7b). Due to this fact this case is not taken into further consideration.

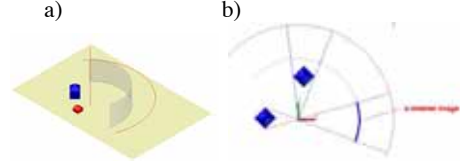


Figure 7: The location of the figure towards the apparatus for receiving a reverse image.

4. FINDING ANALYTICAL RELATIONS ENABLING MAPPING A STRAIGHT LINE ON THE UNROLLED BACKGROUND IN THE INVERSE PANORAMA

The projective relation between points included in any ruling t_F^R , on the unrolled background and the proper points on the ruling t_F^S received in the central projection from the centre S_F onto the projection plane Π means equality between relationships of the double division of homologous fours of points contained appropriately in the rulings: t_F^R , and t_F^S (Fig.3).

For homologous fours of points: $\{^S W, ^S F^{O,S}, ^S P_F, ^S H_F\}$ and $\{W^R, F^{O,SR}, P_F^R, H_F^R\}$ contained in the straight lines: t_F^S and t_F^R , it can be written by the formula below:

$$\frac{^S F^{O,S} P_F}{^S W_{\square,F} ^S P_F} : \frac{^S F^{O,S} H_F}{^S W_{\square,F} ^S H_F} = \frac{F^{O,SR} P_F^R}{W_{\square}^R P_F^R} : \frac{F^{O,SR} H_F^R}{W_{\square}^R H_F^R} \quad (1)$$

$$\Downarrow \frac{^S F^{O,S} P_F}{^S W_F ^S P_F} = \frac{F^{O,SR} P_F^R}{F^{O,SR} H_F^R} \quad (2)$$

Fixing:

- the distance of the point $^S F^{O,S}$ from the centre of base circle \hat{p} by k ,
 - the distance of the point $F^{O,SR}$ from the point $^S P_F^R$ in unreel background by d_0 ,
 - the height of the horizon line by h ,
- and according to Fig.3 the formula (2) can be written as follows :

$$\frac{k \square r}{r \square r_s} = \frac{d_o}{h \square d_o} \quad (3)$$

$$d_o = \frac{h[(k \square r)]}{k \square r_s} \quad (4)$$

In a row series of points contained in the straight line m and series of points contained in the ruling $t_F \square \mathcal{L}$ are projective ones (Fig.4). Moreover, the series $t_F (W_\infty^S, P_F^S, H_F^S, \dots)$ and $t_F^R (W_\infty^R, P_F^R, H_F^R, \dots)$ are congruent.

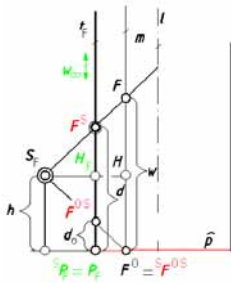


Figure 8: The perspective relation between points contained in the straight line m and the ruling t_F .

It enables establishing the relation below:

$$\frac{|F^O H|}{|F H|} : \frac{|F^O W_\square|}{|F W_\square|} = \frac{|F^{OS} H_F|}{|F^S H_F|} : \frac{|F^{OS} W_\square|}{|F^S W_\square|} \quad (5)$$

$$\frac{|F^O H|}{|F H|} = \frac{|F^{OS} H_F|}{|F^S H_F|} \quad (6)$$

Fixing:

- the distance of the point F from the centre of base plane by w , above formula can be written as follows:

$$\frac{h}{h \square w} = \frac{h \square d_o}{h \square d} \quad (7)$$

$$d = \frac{(w \square h) \square (h \square d_o)}{h} + h \quad (8)$$

Projective relations mentioned earlier help to establish the distance d as the function of the length v of the unrolled panorama background for an angle \square (Fig.3). On the other hand, the representation apparatus is complemented by a rectangular coordinate system $[x,y,z]$ in the way shown in the picture (Fig.9)

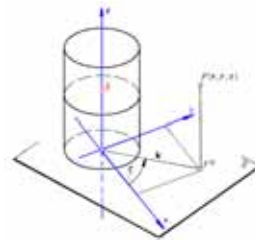


Figure 9: The way of the location of the rectangular coordinate system $[x,y,z]$ for the purpose of analytical presentation of panoramic images.

Every straight line can be written in analytical way according to the rectangular coordinate system $[x,y,z]$ and next it can be transformed to its counterpart according to the rectangular coordinate system $[v,d]$ in unrolled surface. Received analytical algorithms for mapping lines enable drawing inverse panoramic image of any straight line determined by two given points $A(x_a, y_a, z_a)$ and $B(x_b, y_b, z_b)$. Further, it enables drawing images of any spatial figures with computer aid. The algorithms elaborated, have been tested in Mathcad program, however, they can be implemented in any graphical package, which makes drawing more efficient. The figure 11a and the figure 11b show the mapping of the inverse panorama of given classic objects placed the way presented in the picture (see Fig.10).

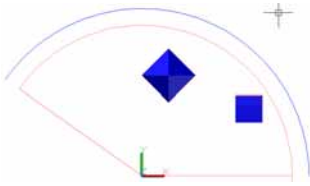
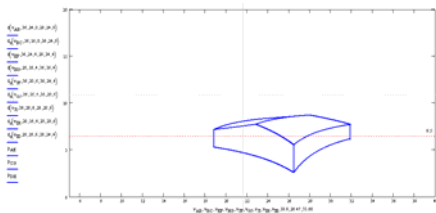


Figure 10: The way of the location of objects towards the apparatus.

a)



b)

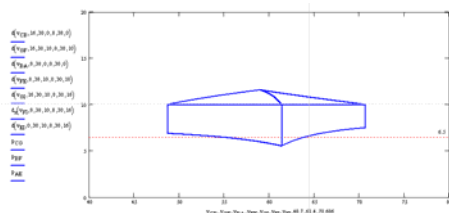


Figure 11: The inverse panorama of the given objects achieved by the computer aid.

5. CONCLUSIONS

In this paper a new way of the direct construction of the inverse panorama is presented. The use of the moving center of the panoramic representation aims at better approximation received projections of the given figures to the real images experienced while watching panorama. The analytical algorithms elaborated and presented in the paper make drawing panoramas very efficient. Therefore, creating inverse panoramas on a flat unrolled background can be

applied in advertising and different artistic presentations.

REFERENCES

- [1] Klette, R., Gimelfarb, G., Kang Wei, S., Huang, F., Scheibe, K., Borner, A., Reulke, R. *A Review on Research and Applications of Cylindrical Panoramas*. Communication and Information Technology Research Technical Report 123, 2003.
- [2] Luhmann, T. *Panorama Photogrammetry for Architectural Applications*. Mapping, ISSN 1131-9100, N° 139, 40-45, 2010.
- [3] Dźwierzynska, J. *Cylindrical Panorama-Two Approaches*. The Journal of Polish Society for Geometry and Engineering Graphics, Vol. 19, 9-14, 2009
- [4] Dźwierzynska, J. *The Use of Two Projective Partly Composed Representation for Constructing Conical Panoramas*. The Journal of Polish Society for Geometry and Engineering Graphics, Vol. 18, 5-10, 2008
- [5] <http://www.panoramafactory.com>
- [6] <http://www.diga.me.uk/360degreeCamera.html>

ABOUT THE AUTHOR

Jolanta Dzwierzynska, Ph.D. is an assistant professor at Department of Architectural Design and Engineering Graphics at Rzeszow University of Technology, Rzeszow, Poland. She can be reached by e-mail: joladz@prz.edu.pl or through postal address: Rzeszow University of Technology, Department of Architectural Design and Engineering Graphics, Poznanska 2, 35-95 Rzeszow, Poland.

RESEARCH AND PRACTICE OF ENGINEERING DESIGN TRAINING WITH A TIGHT INTEGRATION BETWEEN DESIGN AND ENGINEERING GRAPHICS

Zhang Ningrong and Xie Qinghua

PLA University of Science and Technology, China

ABSTRACT: Paying more attention to the application ability of students' engineering practice, the 3D computer aided design training has been discussed with a tight integration between design and engineering graphics. The teaching reform has been put forward on teaching objective, teaching content and teaching method. Some examples of student exercises and course projects show that the new teaching process is more useful to train up the students' innovative consciousness and improve their special capability.

Keywords: Engineering graphics, computer aided design, teaching objective, teaching content, teaching method.

1. INTRODUCTION

As we know, computer aided design has become an especially important technology with benefits such as lower product development costs and a greatly shortened design cycle. Modern CAD systems have adopted 3D solid modeling and many advanced functions, such as finite element analysis (FEA), motion analysis, CNC simulation and machining, and computational fluid dynamics (CFD), and so on. Several solid modeling packages, such as ProE, IDEAS, CATIA, and UG, are academically priced. However, some CAD design courses primarily focused their attention on teaching the CAD packages with no consideration to the theory of orthographic projections and design principles.

Solid models have been introduced as complete and unambiguous computer descriptions of the objects. Having such a formal description available, CAD system can directly perform engineering analysis, manufacture the objects, and, if needed, generate engineering drawings directly from the solid model data base. Therefore, integration between design and engineering graphics becomes a key issue. Initial

graphics integrations occurred in technical reports produced by engineering design teams for integrated projects [1-4]. More recent integration events have been in implementation of graphical analysis techniques [5-8]. We proposed an integrated approach for engineering design training in paper [9]. This paper introduces our experience of teaching the "computer aided design" course, shows some student exercises and course projects. During designing the curriculum, we have installed some electronic classrooms as the autonomous learning platform. The autonomous learning model and the project-driven teaching method have been applied to combine the knowledge and the skill training.

2. COURSE DESCRIPTION

The "computer aided design" course is an introduction to the engineering profession. Whenever possible, our professor discusses or shows the connection between design theory and CAD practice. The 3D CAD practice is the process including solid model design, design analysis and design presentation.

Our teaching contents include engineering conventions and standards, the principles of

descriptive geometry, some CAD software and the engineering analysis methods.

The training outlines have been given including 2D engineering draft, 3D solid model, configuration design, computer aided analysis and design project. The individual units train students to learn the essence of engineering design, the practice of engineering analysis, and to develop computer skills.

2.1 Teaching objectives

The curriculum has the following as central objectives:

1) *Teach the Students How to Read and/or Interpret Technical Presentations:* In the design process, everyone need the skills to (1) read and understand the drawing, (2) to reason about the drawing and (3) to document and output drawing correctly. Engineers are required to know how to interpret 2D drafts as projections of 3D objects, they need to have developed skills of spatial reasoning and they must be able to create proper projections in order to share their designs.

2) *Introduce the Students to Some of the Basic Engineering Tools:* Develop a further understanding of some CAD softwares, and practice presentation skills.

3) *Provide the Students with the Necessary Skills to Perform Effective Problem Solving:* Enable the students to have better spatial analysis skills, develop multiview and pictorial sketching skills, and introduce engineering analysis methods for some solid model applications.

2.2 Teaching context

The training outlines are showed in Table 1. The curriculum has the following as main knowledge:

1) *Engineering Design Graphical Concepts:* They are the principles of descriptive geometry, including the definition of projection, the projections of a point and a straight line, the projections of a plane and a lot of solids, axonometry. Here, we train a double projection method

to study 3D objects through 2D drawings analysis.

2) *Constructions:* Some CAD softwares are introduced, and programs based on computer languages are studied to realize parametric design.

3) *Analysis:* Some qualitative observations about the design efficacy of the model are given. The finite element analysis (FEA) method is introduced.

Table 1: Training outlines.

Topic	Component
Computer Space, 2D Lines	Viewing Computer Space, Drawing 2-D lines, Changing Line Types, Text
2-D Primitives	Drawing 2-D Primitives, Editing 2-D Primitives, 2-D Transformations
2-D Constructions	Tangency Construction, Three-Point Circle, Conic Sections, Splines, Curved Lines
Visualizing Solid Model	Loading Solid Model, Changing 3-D Viewpoint, Hidden Line Removal, Shading Solid Model, Color Hardcopy
Building Solid Model	Base 3-D Primitives, Unary Operations, Boolean Operations, 3-D Transformations, Extrusion Operations, Revolution Operations, 3-D Editing Operations
Analysing Solid Model	Changing Primitives, Redesigning the Model, Mass Properties of a Solid Model, Reverse engineering, Finite Element Analysis of a Solid Model
Prototyping Solid Model	Feature-Based Solid Modeling, Prototyping of a 3-D Solid Model
Projecting Solid Model	Multiview Layout of a Model, Editing Visible Profile Lines, Generating a Drawing

Sectioning Solid Model	Cut Section Operations, Sectioning Conventions, Generating Section Drawing
Dimensioning Projections	Dimensioning Conventions, Generating Dimensioned Engineering Drawing
Design Project	Building, Rendering, and Analyzing Solid Model Assembly, Generating Engineering Drawings

3. TEACHING METHODS

3.1 Autonomous learning

During the implementation of the teaching process, we try our best to foster the students' autonomous ability and to make them become autonomous learners. As a famous Chinese saying goes, "Give a man fish, he will have a meal; teach him to fish, he will have food all his life." As long as students master the ability in autonomous learning, life-long education can be made possible.

We have installed some electronic classrooms as the autonomous learning platform, assisted by electronic laboratory and supported by learning management system. The platform integrated by teaching, learning, practice, test, and judgment, which can provides a completely independent, opening, interactive learning environment for the engineering drawing learners and enthusiasts.

Virtual laboratory can break through the traditional teaching limit, help students better to complete the model of spatial cognition and spatial imagination. As in Figure 1, students can self-operate, rotate or move model in three-dimensional space. Student can observe the model from different points of view, and study different projection on the different projection plane. From the different angle cutting model, students can fully observe the internal structure, and study the shape and projection changes caused by cutting model. Using the

virtual experiment, we develop student's ability of observation, understanding, space imagination and image thinking.

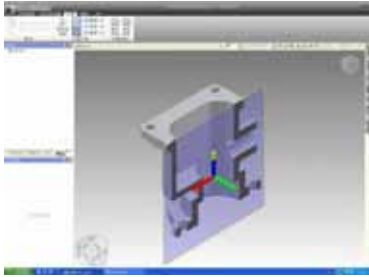


Figure 1: Virtual Experiment.

3.2 Group discussion or team experiential learning

Students should be able to produce technical drawings either by hands or computer generated. When students develop and communicate the design concepts using traditional graphics and computer models, views should be chosen that will best describe the object to be shown. Only the minimum number of views that will completely portray the size and shape of the part should be used. Also, they should be chosen to avoid hidden feature lines whenever possible, as shown in Figure 2 and 3.

Group discussion is often used to determine the view selection on some issues, such as type of spare parts, the main view selection, the other view selection, perspective or oblique view, the whole or partial view. Students can select the multi-view drawing, including front, rear, top, bottom, right, left side views. They also can select the auxiliary and sectional views.

As for team experiential learning, design projects are important components of this course and are carried out in student teams. The teams are encouraged to come up with their own problems to work on, but they could also pick a problem from a given list of ideas or ask the supervisor to suggest a problem. The teams are encouraged to develop their self-confidence,

to enjoy practical and creative work.

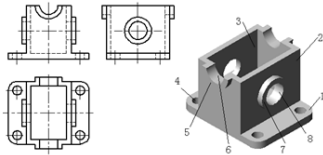


Figure 2: Three-view Drawing.

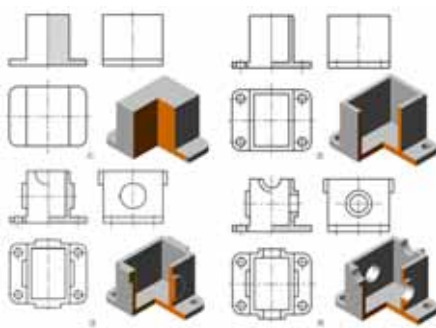


Figure 3: Sectional View.

3.3 Pay attention to configuration design

Creative configuration design of combination solid is a very effective means to improve students' innovative thinking. For example, given the main view, students think space shape and graphical representation of combination solid.

In Figure 4, according to the main view, students may create novel and unique shapes with different surface, such as positive and inclined surface, protuberant and concave surface, planar and curved surface.

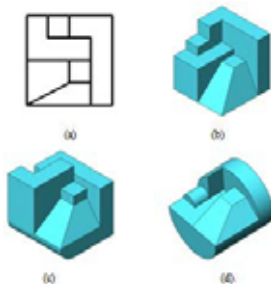


Figure 4: Creative Configuration Design(I).

In Figure 5, according to the main view, students may create novel and unique shapes with different compound mode, such as superposition, excavating, intersecting and synthesizing.

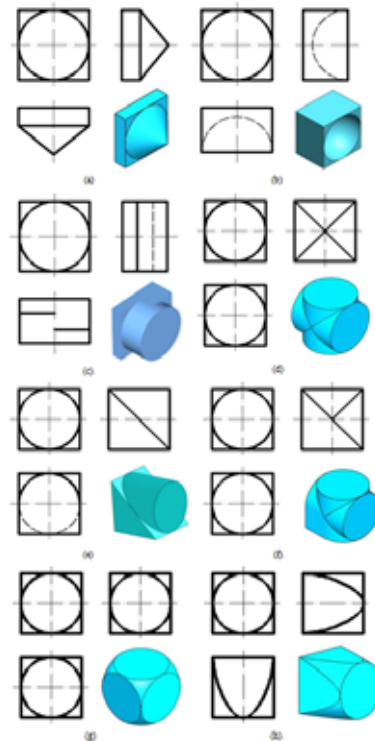


Figure 5: Creative Configuration Design(II).

3.4 Project-driven teaching method

Hands-on activity is a very effective means to improve the application ability of students' engineering practice. In a student-centered classroom, the project-driven teaching method has been applied to combine the knowledge points and the skill training.

For example, we can organize a CAD class following the steps of a) presenting a motivating example, b) introducing a technique to solve the problem, c) solving the problem at

hand, and d) presenting similar problems for the students to solve. The course should be taught in a way that provides students with hands-on experience in applying the techniques taught in class. Students should see an immediate application of the techniques learned.

Students' synthetic practice projects are given in Figure 6 and 7. For instance, students finish the design of ball valve through the following steps of: making sketch map according to part mapping, 3D modeling, virtual assembly, perform interference check, explosion effect, 2D engineering drawing.



Figure 6: Synthetic Practice Project(I).

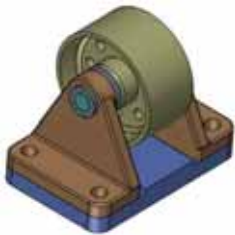


Figure 7: Synthetic Practice Project(II).

Each design project should be ended with a product including a number of components. The whole product should be created as an assembly. Students are encouraged to use advanced functions, including complex curves and surfaces, constraints, visualization functions, cross-sections, different types of views such as detailed views, auxiliary views, revolving views, partial views, and so on, to model the parts, assemblies, and drawings. Each final course project is evaluated by a proposal, a report, and a presentation. We show some exam-

ples of student projects.

3.5 Promote student-teacher communication

To achieve this goal, the course is taught in a classroom in which each student has a computer connected to the Internet. The teacher has a console from which he/she can select any student's computer monitor to be displayed at all computer monitors and projected to the entire class. Using this environment, the teacher assigns a problem to be worked on in class using previously prepared input data made available at the course's web site. Then, the teacher randomly selects a student to show to the class how he/she solved the problem using the technique just learned. This leads to enthusiasm and interactive discussions.

Mapping technology competition has been held once a year. In the spare time and summer vacation, we set up a mapping technology club where tutors and students develop and communicate the new technology of design and engineering graphics.

4. CONCLUSIONS

This paper shows engineering design training should be placed in an integrated environment. In a student-centered classroom, we change teaching methods and the way students learn. Our experiences show that the new teaching process is more useful to train up the students' innovative consciousness and improve their special capability.

REFERENCES

- [1] H. K. Ault, "3D Geometric Modeling for the 21st Century," *Engineering Design Graphics Journal*, 63(2), pp. 33-42, 1999.
- [2] T. Igarashi, S. Matsuoka, H. Tanaka, "Teddy: A sketching interface for 3D freeform design", *SIGGRAPH 99 Conference Proceedings*, pp. 409-416, 1999.
- [3] J. L. Newcomer, E. K. McKell, R. A. Raudebaugh, "Creating a Strong Foundation with Engineering Design Graphics,"

- Engineering Design Graphics Journal, 65(2), pp. 30-42, 2001.
- [4] C. Swann, "Action research and the practice of design," Design Issues, 18(2), pp. 49-61, 2002.
 - [5] T. J. Branoff, N. W. Hartman, E. N. Wiebe, "Constraint-based, solid modeling: What do employers want our students to know?" Engineering Design Graphics Journal, 67(1), 6-11, 2003.
 - [6] Xu Yan, Wang Wei, "Research on engineering graphics teaching combined with 3D solid modeling technology," Journal of Zhejiang Sci. & Tech. University, 22(4), pp. 430-433, 2005.
 - [7] Holdhusen, Mark, "Design of an Engineering Graphics Course for a Pre-Engineering Program", Proceedings of the 2006 ASEE Annual Conference and Exposition, Chicago, Illinois, 2006-1051.
 - [8] K. Fisher, K. Cook, "Improving learning of engineering graphics through a combination of techniques", American Society for Engineering Education, 2007-238.
 - [9] Zhang Ningrong, Yan Li, Chen Yiwang, Chen Baolin, "A systematic approach for engineering design training", Proceedings of 2008 ICCSE – The 3rd International Conference on Computer Science & Education, pp. 1481-1483, 2008.

ABOUT THE AUTHORS

Assoc.-Prof. Dr. Zhang Ningrong, research direction: computer aided design, computer graphics. E-mail: ningrong_zhang@163.com

Assoc.-Prof. Dr. Xie Qinghua, research direction: computer aided design, computer graphics. E-mail: qhuaxie@163.com

RESEARCH OF APPLICATIONS ON SYNCHRONOUS MODELING TECHNOLOGY

Lang WANG, Jing-ying ZHANG, Wei YANG, and Qiong FAN
Beijing Institute of Technology, China

ABSTRACT: Most traditional CAD modeling methods are parametric modeling which sketches are first constructed in 2D planes and surfaces are subsequently formed by sketch curves. All features are placed in the feature tree structure that any feature editing is constrained by the order of creation. However, Synchronous Modeling technology solves the problem. On the basis of Sequence Modeling, the advantages of the Synchronous Modeling have been studied through comparative analysis. Rapid Blue technology is used in Synchronous Modeling. Nonparametric Modeling can be truly and efficiently achieved. Through applying methods, the powerful function and intelligent development direction of Synchronous Modeling are verified. Synchronous Modeling shortens product design time, accelerates product innovation, and improves product design efficiency.

Keywords: Synchronous Modeling, feature tree, CAD modeling, Rapid Blue technology.

1. INTRODUCTION

CAD modeling methods have experienced different stages of development. Wireframe Modeling originates in the late 1960s[1]. It constructs three-dimensional entity with wireframes and polygons[2]. To record topological relations between wireframes and polygons in wireframe modeling, Surface Modeling uses surface table to make the model more stereoscopic. However, the disadvantage of this modeling method is that the model can't be sectioned. With Solid Modeling, the characteristics of solid surface can be controlled by computers [3]. Cutting, stretching, rotating, scanning and other operations of complicated part features can be achieved by using boundary representation method or constructive solid geometry method. It brings a qualitative leap to the three-dimensional modeling technology. The main idea of Parametric Modeling is to explain the shape characteristics of product models through geometric constraints, engineering equations and relationships, so as to design a series of scheme that has the similarity in shape or function[4]. If unreasonable dimension

parameters are given, a feature will interfere with other features, and it leads to the change of topological relations. Parametric Modeling requires full-dimensions constraint. The shape is controlled by dimension constraints. While Variable Modeling technology divides dimension parameters into size constraint and shape constraint[5]. Instead of full-dimensions constraints of all geometric objects, designers can design the shape first and then design the dimensions. Models can be generated under under-constrained and over-constrained conditions, which not only ensure the correctness of design, but also greatly improve the design efficiency. Because of the rapid development of computer graphics, digital design products especially the widely application of 3D technology make product design to a hitherto unknown height[6]. Synchronous Modeling is a new modeling method that is launched by Siemens PLM Software. It is a new breakthrough in digital product development which abandons historical record and changes feature tree to feature set.

In this paper, Synchronous Modeling technology is studied on the basis of traditional CAD modeling methods. The advantages of

Synchronous Modeling technology are obtained through comparative analysis. Through applying methods, the powerful function and intelligent development direction of Synchronous Modeling are verified. The results show that Synchronous Modeling shortens product design time, accelerates product innovation, and improves product design efficiency.

2. APPLICATIONS AND COMPARATIVE ANALYSIS OF SYNCHRONOUS MODELING TECHNOLOGY

2.1 Changing feature tree to feature set



Figure 1 Structure of feature tree and feature set

The difference between feature tree and feature set is shown in Figure 1. For traditional CAD modeling methods, all features are placed in the feature tree structure that any feature editing is constrained by the order of creation. But for Synchronous Modeling, it changes feature tree to feature set, and abandons the concept of historical record completely. Therefore, all features are placed in the same layer in Synchronous Modeling, and feature editing is no longer constrained by the order of creation. No matter for a single part or plurality of components in an assembly, there are no parent-child dependencies between each feature. However, unexpected changes of any feature for traditional CAD modeling methods which are based on feature tree will lead to regeneration of subsequent features. The whole process is a huge time waste when model

structure is very complicated. It requires designers to prepare complicated models in advance. For Synchronous Modeling, design intent can be captured quickly, features can be changed quickly, and the generation of other unrelated features will not be affected. So the model reconstruction in change process can be eliminated. In general, Synchronous Modeling technology can make the design speed increase by 100 times, since there is no need for pre-planning and pre-designing. It Breaks down design barriers based on the inherent structure of historical record.

2.2 Integrating 2D sketches and 3D data together

For traditional CAD modeling methods, sketches are first constructed in 2D planes, and surface is subsequently formed by sketch curves. However, for Synchronous Modeling surface is a directly driven object. User interoperation in Synchronous Modeling integrates 2D sketches and 3D data together. It simplifies the process of CAD operations, and makes 3D as simple to use as 2D. Synchronous Modeling also adopts Rapid Blue technology to break through the limitation of the traditional surface design. It allows designers see the design result immediately. Designers can design any shape of flexible curves through three points which include edit point, control point and contour point. Thereby it maintains the shape of the characteristic and breaks up the traditional model based on surface defects of feature tree structure.

For example, in Figure 2, the wing nut is created with Solid Edge ST5 Synchronous Modeling technology. In this case, the right side of the top cover lacks a plate. To modify this side, users do not have to draw a sketch in a 2D plane, yet can lock a drawing surface on 3D solid, and draw a sketch on the 3D solid directly. The sketches form a closed plane as shown in Figure 3. With Synchronous Modeling technology, the plane can be driven, and the desired dimensions of features can be generated in real time. Synchronous Modeling technology improves the speed and flexibility

of modeling. It gives users the fastest and most flexible experience, and makes it easy to learn for beginners.



Figure 2 No plate on the right side of the wing nut



Figure 3 Sketches form a closed plane on the right

2.3 Using Hand-wheel and Real-time Rules



Figure 4 Hand-wheel and Real-time Rules

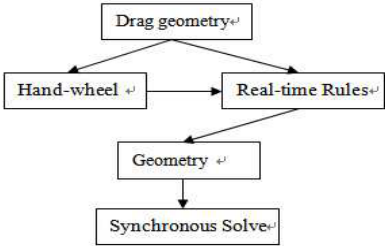


Figure 5 Process of Synchronous Solve

For traditional CAD modeling methods, there is no method to control the model directly. But for Synchronous Modeling technology, the question has been resolved. Hand-wheel and Real-time Rules are show in Figure 4. Taking full advantages of hand-wheel and real-time rules is the key to master Synchronous Modeling technology that will shorts the product development cycle greatly. The hand-wheel is made up of points, axes, ring and plane. With the operation hand-wheel being dragged, rotated and orientated, the model size can be driven quickly. The real-time rules are used to constrain parts when you need to constrain parts during assembling. Real-time rules always guarantee the changes of models, and only features that comply with real-time rules can be changed. The process of Synchronous Solve is shown in Figure 5. In fact, any mouse actions will be detected by the system software. This is the so called real-time rules in Solid Edge ST5. For example, a mining manipulator is shown in Figure 6. If the thickness of the mining manipulator is not suitable for design, click on the face that needs to be modified and then the hand-wheel appears. The distance of extension can be controlled when designers drag the axis of the hand-wheel as shown in Figure 7. It realizes the command of moving features synchronously.

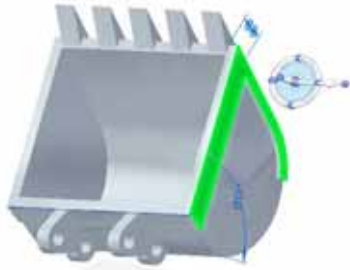


Figure 6 A mining manipulator

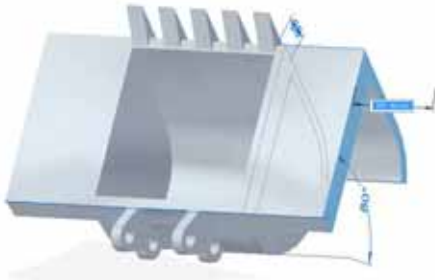


Figure 7 Dragging the hand-wheel

2.4 Adjusting the model easily

For traditional CAD modeling assembly process, the flowchart is simplex as show in Figure 8. If dimensions do not meet the requirements of product design during assembly, users must return to the partial feature and modify the 2D sketches in order to regenerate 3D data, and then go back to assembly interface. Such a separation of modification interface and assembly interface makes the operation process more complicated. However, in Synchronous Modeling, the assembly process as shown in Figure 9, allows users to modify features directly and views of other parts are still visible at the same time. When features are modified, the tree structure of the entire assembly can be navigated. So it can eliminate interference that may exist in the normal sequential modeling, isolate other components quickly, and drive 3D data directly.

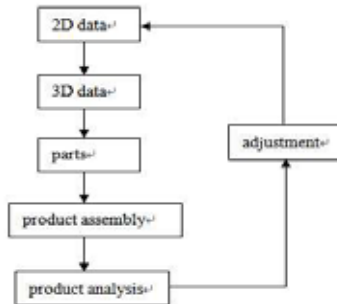


Figure 8 Flowchart of traditional CAD assembly process

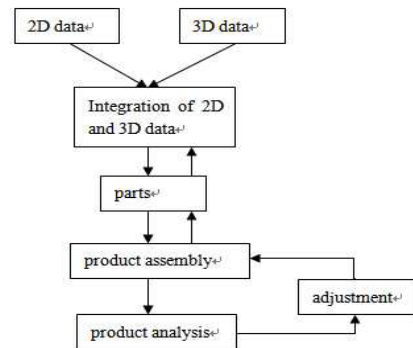


Figure 9 Flowchart of Synchronous Modeling assembly process

As shown in Figure 10, the assembly drawing of a micrometer is drawn with Solid Edge ST5. The diameter of the spindles turns out to be too small during the process of assembly. With Synchronous Modeling technology, the modification of the diameter of the spindles can be done in assembly drawing directly. Likely, the assembly tree can also be operated in the assembly drawing. It's easy to modify features when other parts are selected to be shown or hided appropriately. This operation method simplifies the process of product assembly, especially for the assembly of complicated product. It can shorten the assembly time significantly, improve the speed of product development, and enhance enterprise

productivity.



Figure 10 Assembly drawing of a micrometer

3. CONCLUSION

Synchronous Modeling technology which is based on unconstrained design makes the design more simple and easy. Compared with the traditional modeling techniques, it integrates all the advantages of modeling technologies and breaks down barriers of design based on inherent structure of historical record. For the process of design, Synchronous Modeling technology can be very good to help designers to solve complicated calculating problem, save design time, and get satisfied design results.

ACKNOWLEDGMENTS

The author (L. Wang) is deeply grateful to the Professor Jing-ying ZHANG for her valuable guidance in completing this manuscript. This study was supported by Siemens Product Lifecycle Management Software Inc. Thanks for providing permission of Solid Edge ST5.

REFERENCES

- [1] T. Zhang, Z.Q. Cui, J.H. Hou and Z.H. Zhao. Modeling Technologies and Their Features in CAD Evolution. *Mechanical Engineering & Automation*, (6): 196-200, 12 2011.
- [2] Q.C. Ma, L.H. Liang, S.H. Lie and Y.N. Luo. The Research of UG Modeling Method. *Machine Design and Research*, 18(1): 40-42, 1 2002.

- [3] S.R. Yang, S. Hu and J.J. Li. CAD data exchange and interoperability. *COMPUTER ENGINEERING AND DESIGN*, 31(7): 1581-1584, 6 2002.
- [4] X.J. Chang and J.P. Ren. Both Modelling Technologies of Modern CAD System. *MECHANICAL MANAGEMENT AND DEVELOPMENT*, (1):85-86, 2 2007.
- [5] D.F. Liu, J. Tao, B. Wang and J. Wang. Research on the Face Finder and Behavior Decoding Technology in Synchronous Modeling. *Advanced Materials Research*, 308:67-71, 2011.
- [6] J. Alan. A beginner's guide to Synchronous Technology. *Manufacturers' Monthly*, 2008.

ABOUT THE AUTHORS

1. Lang WANG is now pursuing a master's degree at Beijing Institute of Technology. His current research interest includes mechanical design, mechanical simulation, and computer graphics.
2. Jing-ying ZHANG is a Professor of Engineering Graphics at Beijing Institute of Technology, China. She received her PhD in Machinery Manufacturing and Automation from Beijing Institute of Technology, China. Her current research focuses on advanced cutting technology, Mechanical Computer Aided Design and computer graphics.
3. Wei YANG is a Lecturer of Engineering Graphics at Beijing Institute of Technology, Beijing, China. She got her Master Degree of Engineering in Machinery Manufacturing from Beijing Institute of Technology, China. Her current research focuses on digital design of mechanical and computer graphics.
4. Qiong FAN is now pursuing a master's degree at Beijing Institute of Technology. She is majoring in mechanical engineering. Her current research focuses on ultrasonic non-destructive testing and three-dimensional modeling.

A RESEARCH OF MULTIMEDIA TEACHING MATERIALS FOR 3-DIMENSION CAD EDUCATION

Hiroki TOMINAGA¹, Naoki IIDA¹, Masayuki MORI² and Kenichi HIROSE¹

¹College of Industrial Technology, Japan ²Osaka University, Japan

ABSTRACT: In this paper, we report the educational results that used multimedia teaching materials for graphic science education in the Department of Information Engineering, at the College of Industrial Technology. The class of “CAD” is aiming to develop “spatial ability” through lectures and exercises using Autodesk Inventor that is 3-Dimension CAD software for graphic science education. The students are learning operation methods of Autodesk Inventor, from easy level to difficult level exercise problems. We made multimedia teaching materials for students to review lectures and exercise problems. The students can select multimedia teaching materials, according to the level of understandings. According to the questionnaire survey of the evaluation by the students, we consider about the educational effects and problems.

Keywords: 3-Dimension CAD, e-Learning, Multimedia Teaching Materials

1. INTRODUCTION

Various improvements to educational methods have been made using ICT (Information and Communication Technology), in the Department of Information Engineering, at the College of Industrial Technology. PDF (Portable Document Format) [1] files of teachers’ lecture materials and other teaching materials are sent to an LMS (Learning Management System) and lecture videos which are records of all lectures by teachers are sent to a video server as teaching material, and these are applied to practical subjects that use computers, such as programming education or visual information processing education and so on. Lectures using videos in particular, can help students concentrate on the lessons, confirming their learning effectiveness [2-4]. The class of “CAD” is aiming to develop “spatial ability” through lectures and exercises using Autodesk Inventor that is 3-Dimension CAD software for graphic science education, in the Department of Information Engineering. From practice problems using simple shapes through constructive exercise problems, students master software operation methods and

develop spatial abilities. Experience is necessary to master software operation methods, and it is important to give students many practice problems or exercise problems. This class is also taught using LMS and a video server to deliver teaching materials, but gaps between the scholastic levels of the students have widened, and it has become difficult to bring them back to the same scholastic level with only one set of teaching materials. So multimedia teaching materials were prepared, and partly applied to teaching, confirming their educational effectiveness [5].

This research reconstructed these teaching materials for use in “Mastering Software Operation Methods and Nurturing Spatial Ability”. The practice of this education is reported at the same time as the effectiveness of and problems with this education are considered based on the results of a questionnaire survey of students’ evaluations of lessons.

2. MULTIMEDIA TEACHING MATERIALS

The following are multimedia teaching materials we made.

2.1 Lecture materials with PDF

These PDF files are the presentation slides of teacher. Mainly students use these PDF files to check the contents of lecture and exercise problem. Lecture contents, practice problems or exercise problems etc. are shown, and students confirm these using the PDF files presentation slides. Figure 1 shows Lecture materials with PDF.



Figure 1: Lecture materials with PDF.

2.2 Operation method materials with PDF

These PDF files are the presentation slides of the screenshots which include every operation of Autodesk Inventor. It contains information ranging from how to start the software to basic operation methods and practice problem preparation methods, so students can check operation methods and prepare practice problems by referring to this document. However, as the problems advance, operations already mastered are omitted. Figure 2 shows Operation method materials with PDF.

2.3 Lecture videos

These videos are recorded lectures that used the presentation slides of teacher. Detailed explanations are added by reading in the teacher's voice or inserting comments to the lecture contents and to the practice problems or exercise problems. The overall lecture by a teacher is recorded, so students absent from a lecture use it to confirm the lecture contents. Because there are detailed explanations by teachers, students who could not understand simply using the PDF

file materials use it to prepare practice problems and to confirm or review the lecture contents. Figure 3 shows Lecture videos.

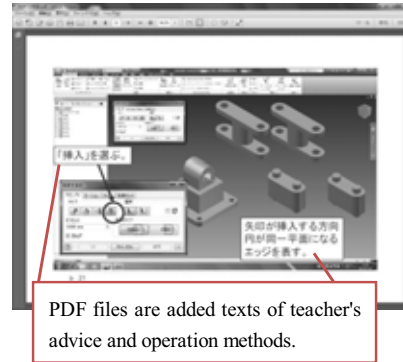


Figure 2: Operation method materials with PDF.

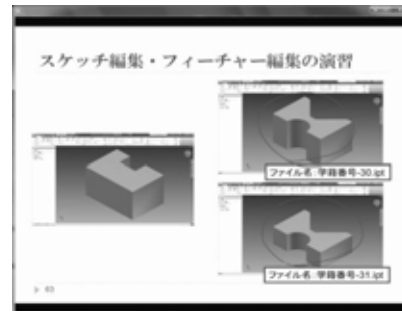


Figure 3: Lecture videos.

2.4 Operation method videos

The videos were prepared as a manual by adding detailed explanations of a model operation method for teachers to operate Autodesk Inventor 11 in the voice of the teacher. All details of operation are recorded in the videos, so students who were unable to understand this using only the PDF file materials can use the videos to confirm the operation method. Also, after students submit answers to exercise problems or problems for mid-term tests, teachers present model answer videos, so students can use them to confirm and review operation methods as

end-of-term test measures. The lengths of the videos differ by problem, and sample solutions to problems are each completed by one video. Figure 4 shows Operation method videos.



Figure 4: Operation method videos.

We provided these PDF files by LMS (Learning Management System), and uploaded these videos in the video server. The students can select multimedia teaching materials which they access, according to the level of understandings.

3. EDUCATION PRACTICE RESULTS

Lectures were given in 2012 and 2013 using multimedia teaching materials. The class of “CAD” is 90 minutes × 2 frame lectures done 15 times. The two PDF files, which are multimedia teaching materials, were uploaded appropriately for each lecture, so it was possible to download them at any time from the LMS. Lecture videos recorded teachers’ lectures in real time, so they were uploaded to a video server within a few days after the end of the lectures. Operation method videos which are prepared in advance can be watched immediately, but videos recorded based on a judgment that explanations made in answer to questions asked during the lecture are necessary as teaching material were uploaded like the lecture videos on a later day. The flow of the lecture was

simple explanations such as a review of the previous lecture and practice problems or exercise problems based on the teacher’s presentation slides. After the explanations, during exercise time, students selected lesson material suited to each person to do exercises, and filled out lesson evaluation questionnaires when the lecture had been completed. A total of 37 students took the course in 2012 and 39 students took it in 2013.

Table 1 shows the lesson plan in 2011, while Table 2 shows the lesson plans in 2012 and 2013. Problems in 2012 and 2013 are based on 2011, with new problems added. Table 3 shows the number of problems in each year.

Table 1: Lesson Plans in 2011.

Times	Outline
1	Guidance
2	CAD systems and 2-dimensional and 3-dimensional CAD
3	
4	Foundations of graphic science and perception of a figure
5	
6	Mid-term test (graphic science)
7	Basic operation of Inventor and Exercise Problems
8	
9	
10	
11	
12	
13	Mid-term test (modeling)
14	Nurturing spatial ability centered on modeling
15	Final problem (modeling)

Table 2: Lesson Plans in 2012 and 2013.

Times	Outline
1	Guidance
2	CAD systems and 2-dimensional and 3-dimensional CAD
3	
4	Foundations of graphic science and perception of a figure
5	
6	Mid-term test (graphic science), Basic operation of Inventor and Exercise Problems
7	Basic operation of Inventor and Exercise Problems
8	
9	
10	
11	Mid-term test (modeling)
12	Nurturing spatial ability centered on modeling
13	
14	Final problem (modeling)

Table 3: Number of problems each year.

Times	Number of problems		
	2011	2012	2013
1	-	-	-
2	5	6	6
3	5	5	5
4	1	1	1
5	1	2	3
6	-	5	5
7	5	22	22
8	12	6	6
9	10	18	18
10	1	2	2
11	5	-	-
12	7	2	2
13	-	1	1
14	4	1	2
15	1	-	-
Sum	57	71	73

In 2012 and 2013, the number of frames of “Basic operation of Inventor and Exercise Problems” was lower than in 2011. This is because in 2012 and 2013, it was possible to shorten the time spent explaining details such as basic operations, practice problem preparation methods etc. by using teaching materials prepared in advance, instead of by giving the teacher’s explanations in class as in 2011. The number of frames is low, but the students’ exercise time was long, so it was possible to give many problems in the “Basic operation of Inventor and Exercise Problems” period, concentrating exercise content. The added problem was preparing a model from an isometric drawing, and by repeating the basic operation during this process, students learned basic operation. Figure 5 shows an example of an exercise problem. “Nurturing spatial ability centered on modeling” and “Final problem (modeling)” each increased it by 1 frame. The problem, “Nurturing spatial ability centered on modeling”, is an exercise: creating by combining multiple prepared models. It is possible to confirm the collision detection of models, so it is possible to make a model which functions. Figure 6 shows a sample work of a Web camera, and Figure 7 shows a sample work of a robot.

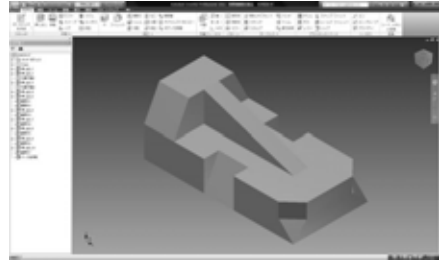


Figure 5: example of an exercise problem.

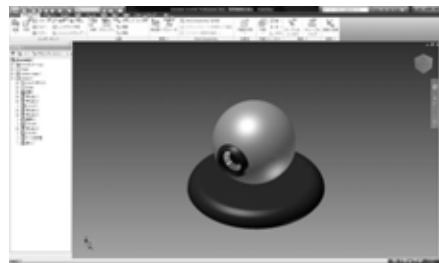
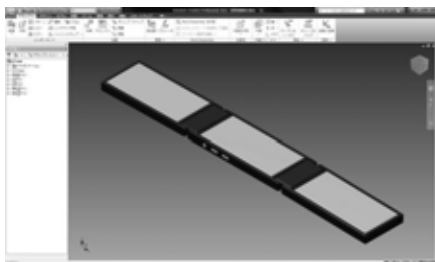


Figure 6: sample work of a Web camera.

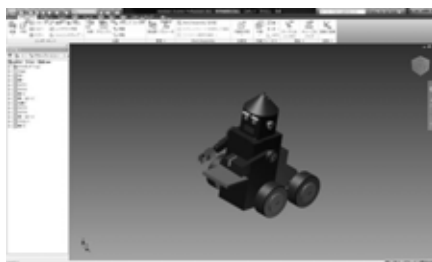


Figure 7: sample work of a robot.

“Final problem (modeling)” in 2013 applied modeling of an original cell phone, which was the final problem in 2011, and modeling of an original robot, which was the final problem in 2012. Figure 8 (a) shows a sample work of an original cell phone and Figure 8 (b) shows a sample work of an original robot.



(a) sample work of an original cell phone.



(b) sample work of an original robot.

Figure 8: Final problem (modeling)

Table 4: Evaluation standards in the lesson evaluation questionnaire.

Item	Evaluation contents			
Difficulty	Easy	Somewhat easy	Somewhat difficult	Difficult
	1	2	3	4
Number of problems	Few	Somewhat few	Somewhat many	Many
	1	2	3	4

Presumably, the reason why it was possible to apply two problems, is that through many problems, the method of operating Inventor was firmly established, even applications centered on the foundations, shortening the time spent preparing each problem. Table 4 shows the evaluation standards in the lesson evaluation questionnaire.

The question, “What score do you give to your own efforts in the class out of a perfect score of 100?” was posed with a free answer space provided. Figure 9 shows the results of the questionnaire concerning difficulty and number of problems (average values) in 2013.

Difficulty and number of problems both earned good evaluations at 7 times, 9 times, and 14 times, but at each number of times, many models are prepared, so this result was obtained. At 10 times and 12 times, there was a gap between number of problems and difficulty, but this is thought to be because problems that make the students consider the question, “by what pro-

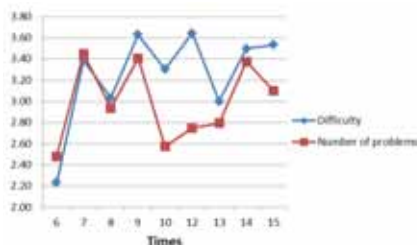


Figure 9: results of the questionnaire concerning difficulty and number of problems (average values) in 2013.

cedure, how is it completed?” were given. Table 5 shows the results of the questionnaire survey (average values) concerning the students’ efforts in the class.

Table 5: Questionnaire results.
(What score do you give to your own efforts in the class out of a perfect score of 100?)

Times	6	7	8	9	10	12	13	14	15
75 or higher	58.3%	75.0%	71.0%	66.7%	75.8%	50.0%	80.0%	56.5%	65.5%
Over 50 and less than 75	41.7%	25.0%	25.8%	29.6%	21.2%	50.0%	20.0%	34.8%	34.5%
Over 25 and less than 50	0.0%	0.0%	3.2%	3.7%	0.0%	0.0%	0.0%	8.7%	0.0%
Below 25	0.0%	0.0%	0.0%	0.0%	3.0%	0.0%	0.0%	0.0%	0.0%
Average degree of efforts	80.46	80.68	78.61	78.59	80.55	77.14	82	76.3	81.38

The score was lower than 50 at 8, 9, 10, and 14 times, but this is because some students were absent from the previous lesson and had not reviewed. At 12 times and 14 times, the average score fell, but the reason is thought to be that each time, the only themes “Web Camera” and “Robot” were given, and they had to consider these beginning with design. The rise at 13 and 15 times is thought to be a result of the fact that it could be concentrated at completion of the work from the design drawings. The following are excerpt from students’ opinions given in the free answer space.

- While doing problems, I learned about CAD, so I have changed my mind and think the number of problems is just right.
- With CAD, it was difficult to clarify space, but I was happy when the work was completed.
- There were many problems, so it was difficult, but it was worth the trouble.

When exercises started, many said they were too difficult or there were too many problems, but through the problems, somewhat good evaluations of “Mastering Software Operation Methods and Nurturing Spatial Ability” were obtained.

4. CONCLUSIONS

This report shows that the results of the student questionnaire verified the effectiveness of education with multimedia teaching materials during CAD lectures in 2013. Providing explanations of the software operation method by

preparing PDF and video documents permitted students to select the lesson material to perform exercises matched to their degree of understanding. Also, long exercise time and many problems permitted the students to repeatedly consider the question, “by what procedure, how is it completed?” to master the operation method. But, in cases where a student who was absent from a lecture did not do a problem until the next time, even if the student confirmed the content of the problem, a gap in scholastic level appeared. It is now possible to access a video server only on campus, so we must improve it so it can be accessed from outside the campus, to encourage students to study.

In this report, it was multimedia teaching material used for “Mastering Software Operation Methods and Nurturing Spatial Ability”, but it is presumed that if there is teaching material used to acquire knowledge of graphic science, it will be more effective. We want to also study the expansion to flipped teaching: letting the students see multimedia teaching materials before classes begins.

REFERENCES

- [1] Adobe Systems Incorporated. PDF. <http://www.adobe.com/devnet/pdf.html>. Accessed 6 2014
- [2] Kiyotsugu SATOH and Toyohisa KANEKO. On the Study Effect of e-Learning in the Engineering Experiments. *BULLETIN OF COLLEGE OF INDUSTRIAL TECHNOLOGY*, Volume 40: 87-92, 3 2006.

- [3] Kenichi HIROSE, Masayuki MORI and Kiyotsugu SATOH. A Study of Interactive and Visual Learning Material in the Computer Graphics Education. *BULLETIN OF COLLEGE OF INDUSTRIAL TECHNOLOGY*, Volume 41: 93-98, 3 2007.
- [4] Kenichi HIROSE, Hiroki TOMINAGA, Masayuki MORI, Toyohisa KANEKO, Naoki IIDA and Kiyotsugu SATOH. A Study of Visual Learning Material for Practical Programming Education. *BULLETIN OF COLLEGE OF INDUSTRIAL TECHNOLOGY*, Volume 42: 63-68, 3 2008.
- [5] Naoki IIDA and Hiroki TOMINAGA. Development of Teaching Material of Video Contents Depending on the Understanding of Students for the Education of 3-Dimensional CAD. *BULLETIN OF COLLEGE OF INDUSTRIAL TECHNOLOGY*, Volume 43: 49-56, 3 2009.

ABOUT THE AUTHORS

1. Hiroki TOMINAGA is an Assistant in the Learning Support Center, College of Industrial Technology. His research interests are e-Learning. He can be reached by e-mail: tominaga@cit.sangitan.ac.jp, by FAX:

+81-6-6431-7992, by postal address: 1-27-1 Nishikoya / Amagasaki / Hyogo / JAPAN.

2. Naoki IIDA is an Associate Professor in the department of Information Engineering, College of Industrial Technology. His research interests are e-Learning. He can be reached by e-mail: iida@cit.sangitan.ac.jp, by FAX: +81-6-6431-7992, by postal address: 1-27-1 Nishikoya / Amagasaki / Hyogo / JAPAN.

3. Masayuki MORI is an Assistant Professor in the Cybermedia Center, Osaka University. His research interests are Interactive Three-dimensional Computer Graphics. He can be reached by e-mail: mori-m@cmc.osaka-u.ac.jp, by postal address: 4F 1-32 Machikaneyama / Toyonaka / Osaka / JAPAN.

4. Kenichi HIROSE is a Professor in the department of Information Engineering, College of Industrial Technology. His research interests are Image Processing, Computer Graphics and Human Interface. He can be reached by e-mail: hirose@cit.sangitan.ac.jp, by FAX: +81-6-6431-7992, by postal address: 1-27-1 Nishikoya / Amagasaki / Hyogo / JAPAN.

THE RESEARCH ON GRAPH, GRAPHICS AND GRAPHICS SCIENCE

Baoling Han¹ and Yuanjun He² and Bingshu Tong³ and Xiao Luo¹ and Jie Yang⁴

¹Beijing Institute of Technology, China ²Shanghai Jiao Tong University, China

³Tsinghua University, China ⁴China Graphics Society, China

ABSTRACT: Like words and digits, graph has an irreplaceable effect in social progress, economic construction and science-technology development. But it still exist the problems of how to define and recognize the role for graph and graphics from the view of history and present-day. At present, some disciplines, such as Engineering Graphics, Computer Graphics, Computer Images, etc., have their own disciplinary position and independent development mode. These disciplines may have some dispersion and limitation, because they lack links to graph and graphics. This thesis shows the nature of graph from the view of shapes (or geometry), reveals the internal relationship between shape and graph and gives the scientific description for shape and graph. And also proposes the concept of “Great Graphics” to integrate related graphic science and technology scattered in other disciplines^[1], in order to take research on basic theory and approach and application of graphics.

Keywords: Graph, Graphics, Graphics Science.

1. INSTALLATION

As we know, graphics is one kind of subject for which used to do some research on the relationship between the "graph" and the "shape". With the development of computer technology, and the outbreak of the digital revolution, the graphics technique becomes one of the key technologies in this world, while the developing emphasis of graphics has transferred to the relationship between the "figure" and the "shape", which means the research of graphics focus on how to describe, storage and transfer the shape and the graph in digital form. The research of graphics has been carried forward not only on computer graphics, computer arts, but also contains digital design and manufacturing (CAD/CAM), modeling and simulation technology and so on. The traditional graphics has turned into a modern graphics discipline and adapt to the global digital information age. The research content of graphics contains some different development directions: visualization in scientific computing, virtual reality system, realistic graphics, and fractal picture, computer

animation, image fusion technology and so on. This thesis focuses on the theory and application of graph expression, generation, processing and communication during the procession of deduction from shape to graph and reverse, in order to take research on basic theory and approach and application of graphics. This thesis focuses on the theory and application of graph expression, generation, processing and communication during the procession of deduction from shape to graph and reverse, in order to take research on basic theory and approach and application of graphics^[1]. The graphics science and technology has becomes a new kinds of science and technology studies, an interdisciplinary of speculative geometry, computer science and engineering design graphics.

2. GRAPH, PATTERN, FIGURE AND IMAGE

2.1 Graph, shape and pattern

As we know, graph is our common language, graph is the target and the visual representation of shape, while shape is the source of graph.

This means that shape is the representation of objective or virtual world, while graph is the exhibition of shape in screen^[2]. We use graph to describe this world, for example, we use natural image to reflect the world, use representation graph to reveal the world, use creative image to imagine the world. The natural image contains photograph (Figure 1), remote sensing image (Figure 2), cloud chart, maritime map, CT image (Figure 3), electrocardiogram (Figure 4), defectogram and so on, these graph show the objective things with the aid of tools, just like we observe the scene by eyes. Creative image contains painting (Figure 5), conception (Figure 6) and so on, which is the crystallization of human thinking and imagination. Representation graph contain geometric graph, projection drawing, axonometric drawing, analysis chart, hand drawing and so on, this kind of graph is formed through specific rules.

Natural image:



Figure 1 Photograph



Figure 2 Remote sensing image

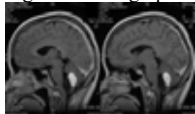


Figure 3 CT image



Figure 4 Electrocardiogram

Creative image:



Figure 5 Painting



Figure 6 Conception

As a language of science Technology, pattern usually use to convey design idea and exchange technology, which means build a bridge between design and implementation in engineering field. All in all, pattern complements human language, beyond that, pattern also is the

concrete reflection of human intelligence and language in the higher developmental stage.

Engineering drawing is used to express the requirement for engineering technology, for example, according to the projection principle and standard code, we use engineering drawing to express the requirement about the shape, size and technical requirements of the project object. Engineering drawing is the expression vector of engineering and product information, and it is a language that used for expression and transformation in engineering field, in addition to these, with a view to its great application such as mechanical drawing, construction drawing, piping design and so on, it is also a kind of important technical document in modern production.

2.2 Figure and image

Figure is a kind of vector diagram which comprised by exterior contour lines, such as straight line, circle, rectangle, curve, chart and others which drawn by hand or computer. The important element of figure is the topological relation between graphic elements, for example, connected relation, joint, tangent relationship, parallel relationship and so on. In consequence, as a two-dimensional structure, the main mission of generating figure is to decide the relationship between the geometrical elements which form a figure, a typical science of processing figure is “computer graphics”.

Comparing with figure, image comes in the form of lattice diagram, which means it stresses the overall form, and it is used to record the information such as gray scale and color of points in figures. In display respect, image can express the object in definite resolution ratio and then show the color information of every point in digital way, by this way, it can quick display on the screen, a typical science of processing image is “computer image processing”.

2.3 Graph, shape and pattern

The difference between figure and image in computer science is as follow: figure usually means the graph which drawn by computer, such as straight line, circle, circle arc, freeform curve, diagram and so on. While image is another kind of graph, both the actual scene images that can be caught by input equipment or any picture that stored in digital form are representative examples.

On the step further, take the advantage of flexible and creative, image is composed of pixels in a particular order, not only can be used to express the real-world scenarios, but also can express the details of the complicated drawing.

We use figure to record the algorithm and characteristic of sprouting graph which gives it another name “vector diagram”. Its characteristics such as easy to translate, scale, rotate, twist will mainly be used for line drawing, engineering drawing, artistic calligraphy and so on.

3. GRAPHICS AND GRAPHICS SYSTEM

3.1 Graphics definition

As mentioned, the study object of graphics is graph and we use graphics to simulate the real world, construct the virtual world. Graphics refers to the science that its center is graph, the research about it focus on the expression, generation, processing and translation during the process of shape deducting to graph and vice versa.

In the days before computer, the generation of graph mainly relies on hand-painted, and because of this, the correlation theory and method of manual drawing has been seen as the main content of graphics. For example, the research on the engineering drawing theory and method comes into “engineering graphics”, which means its theoretical basis is descriptive geometry, projective geometry and so on. Currently, engineering mathematics, engineering

mechanics and engineering graphics are treated as the basics of engineering course [3, 4].

We combine graphical theory, computer technology, engineering graphics applied mathematics, physics and other subject to create the new “computer graphics”, in addition to this, there are other new subjects such as “computer-aided design (CAD)”, “computer aided geometric design (CAGD)” and so on, besides of this, it also create some new subsience and application areas, for example, visualization in scientific computing, virtual reality, man-machine interaction, computer animation and so on.

3.2 Graphics system and branch

The emergence of increasing new cross technology subjects along with the development of science and technology. At the time when computers are widely spread, the application of figure/image is manifested in every domain of social life and production. In the middle of the 20th century, computer and automatic drafting machine came out, and it has given birth to the interactive graphics. The first mentioned of “Computer Graphics” in the PhD thesis “Sketchpad: A man-machine graphical communication system” of Sutherland from Massachusetts Institute of Technology in 1962 proved that interactive computer graphics is a feasible and useful research area, and his work establishes the independent state of computer graphics as a new subsience. In 1960s, the newborn computer aided geometric design lays the theoretical foundation of the complex calculation of curved surfaces. The researches above facilitate the integration of graphics, mathematics and computer technology and form a new discipline, which name is computer graphics; the appearing of this discipline marks the modern graphics technology era. The rapid development of CAD remaining with the de-

velopment of computer graphics, and all the changes signals the new era of paperless. Later, there are two main branches come out, entity modeling technology and realistic graphics generation technology. The major research of geometric modeling is simulating “shape”, while the major research of realistic graphics generation technology is output of “graph”, and all of the researches above have been unified together in computer graphics harmoniously. In the following research of image processing directly, the limit between modeling and plotted output gradually blurred, and what happened also help computer graphics coming into the world. Now the research of figure and image can often be much the same thing. Therefore, Prof. Yuanjun He and Prof. Bingshu Tong proposed a new concept which name is “Great Graphics” and began to study graphics science and technology.

3.3 Graphics subject system

Several subjects such as speculative geometry, numerical mathematics, computer science and engineering design closely related to graphics, therefore these can be seen as the interdisciplinary subjects of graphics. According to the research on the definition and nature analysis of shape, graph and graphics, Prof. Yuanjun He and Prof. Bingshu Tong proposes a new three layers graphics subject system framework which is composed of graphics based layer, application support layer and graphics application layer. The framework is shown in Figure 7 and Figure 8.

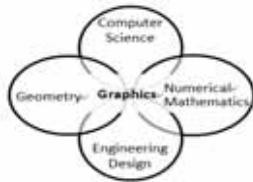


Figure 7 Interdisciplinary subjects with graphics

As can be seen from the above diagram, graphics based layer contains graphics public basic, computation basis and graphics theory. Application support layer contains graphical based software, parts library and product information modeling. Graphics application layer: As we know, graph has been used in our lives widely, such as Engineering and product design, drawing, graphics design, graph creativity, information visualization, geography and information system and other application fields which based on graphics. Graphics standard: As the communication language of scientific, engineering, artistic and so on, and a kind of tool that can transfer the design and knowledge, graphics should follow specific standard and criterion, which contains the graph production standard, graph interchange format (such as STEP, DXF), image storage and interchange format (for example JPG, BMP and so on).

4. THE DEVELOPMENT OF GRAPHICS RESEARCH

Although graph and graphics all over many fields, there are not so many people begin to study it as it is an independent discipline, all we have done in theoretical aspect just driven by application demand and the research result belongs to the related disciplines. Now the research about graphics is scattered in different areas such as machinery, computer, information communication, electronic technique, mapping and so on, and the current situation reflects that graphics is an interdisciplinary subject. So we think the development of graphics research can be broken through in following directions.

4.1 Information fusion of figure and image

There are so many research directions explain that the application of interdisciplinary subject demands for the new algorithms and theory, for example, image characteristics extraction, image acquisition and tracking in special environment, all of these lay the foundation of multi-disciplinary integration.

4.2 Augmented reality

Along with the deeply research into augmented reality and virtual reality, we can use voice, gesture, touch to perceive the world and learn about the nature.

4.3 Information visualization

In order to meet the demand of information technology, big data handling techniques make information visualization come true. Besides, big data, cloud computing and the Internet of Things become the main component of intelligent system.

4.4 3D printing

With the development of the new material and design modeling technology, 3D printing has become more mature. This technology challenges the traditional product development mode because it can shorten the design cycle and which means the competition of the concept product becomes more and more intensely.

5. CONCLUSIONS

With the great impact of the new technology, new concept and new demand, the research on graphics will develop towards the high-technology direction. And the new technology such as information fusion of figure and image, augmented reality, information visualization, computer arts, animation production and so on will be widely used in engineering design, aerospace engineering, daily life and other fields. We hope more and more people can participate into the research on graphics, this work can change our life.

ACKNOWLEDGMENTS

I would like to express my gratitude to all those who helped me during the writing of this thesis. First of all, I would like to extend my sincere gratitude to all collaborators, for their instructive advice and useful suggestions on this thesis. I'm deeply grateful of their help in the completion of this thesis. Special thanks should go to my family and friends who have put considerable time and effort into their comments on the draft.

REFERENCES

- [1] Zhang Tuo. A tentative study on the development of the graphics discipline in the domestic and overseas. *Journal of Xiamen Educational College*, 2011.13 (3):36-40.
- [2] Bingshu Tong. Cognition to the Discipline of Engineering Graphics and the Discipline of Graphics[J]. *Journal of Engineering Graphics*, 2010.31 (6):1-6.
- [3] YuMing Ding. On the contents and branches of the discipline of engineering graphics[J]. *Journal of Engineering Graphics*, 1998 (1):84-89.
- [4] China Association for science and technology, China Graphics Society. Report on Advances in Graphics. China Science and Technology Press. 2014.4.

ABOUT THE AUTHORS

1. Baoling Han, doctor of engineering, professor and Ph.D. supervisor in School of Mechanical Engineering and School of Optoelectronics at Beijing Institute of Technology. At present, she is a member of ISGG, executive director of CGS, director of the International Liaison Committee for CGS, director of University Creative Education Branch of Chinese Invention Association, and senior member of Chinese Mechanical Engineering Society.

Her e-mail and postal address is as follows: hanbl@bit.edu.cn. School of Mechanical Engineering, Beijing Institute of Technology, 5 South Zhongguancun Street, Haidian District, Beijing 100081, China.

2. Yuanjun He, Shanghai Jiao Tong University, China. His e-mail address: luox-since@163.com.

3. Bingshu Tong, Tsinghua University. His e-mail address: 2426447470@qq.com.

4. Xiao Luo, Beijing Institute of Technology. Her e-mail address: luox@bit.edu.cn.

5. Jie Yang, China Graphics Society, China. Her e-mail address: 317834051@qq.com.

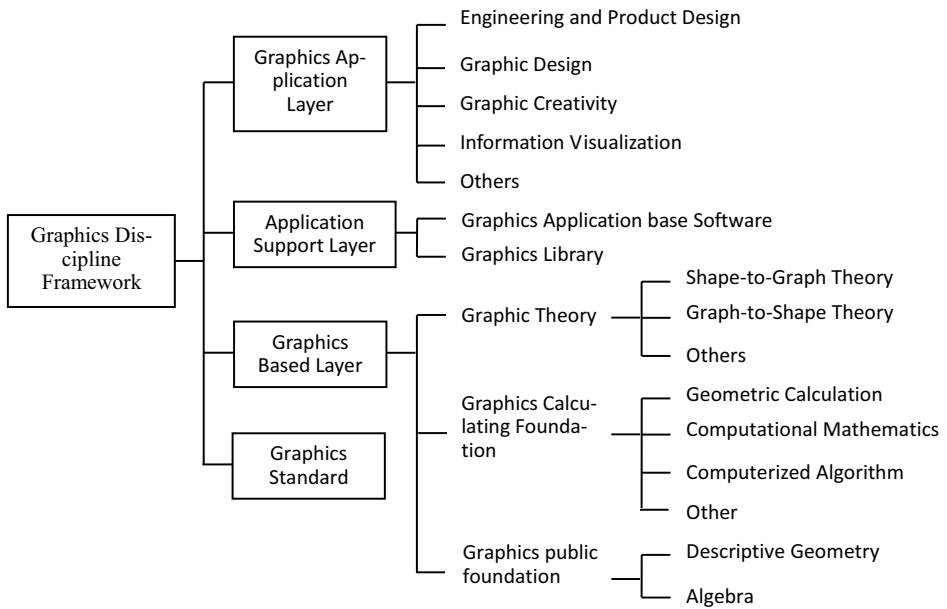


Figure 8 Graphics Discipline Framework

RESEARCH ON THE TEST SHEET ORGANIZATION FOR ENGI-NEERING GRAPHICS

Lijie Guan, Huilai Liu, Chunhua Wang and Feng GUO
Northeast Petroleum University, China

ABSTRACT: The main functions and characteristics of engineering graphics test bank are discussed. The various organizations of test paper in the test bank are described in details. A management system of auto-organizing test paper designed with visual interface is put forward. Users can organize a certain test paper and preview, modify or delete the default test paper in the system. Users can also set types, chapters, quantity, difficulty and score of questions to create a new test paper. The management system is convenient and reasonable in automatic generation of test paper, thus is of practical value.

Keywords: Engineering graphics, test bank, organizing test paper, option of organizing test Paper.

1. INTRODUCTION

Examination is mainly to encourage the students to study, and to measure the effect of teaching. Standardizing the entire procedure of examination and minimize the impact of human factor on examination to achieve the objectivity and scientificity of examination is not only the pursuits goal of teaching, but also the foundation to achieve the separation of teaching and examination and to further implement credit system[1][2]. Thus the following are required: firstly, an effective management model; secondly, a scientific and practicable test bank management system.

Whether a test bank is successful or not is judged in terms of the following two aspects [1]. Firstly, the test papers should reflect knowledge level and ability of students accurately, reliably, effectively and prospectively, and can distinguish abilities of individuals. To achieve practicability of the test papers, the test questions must cover the widest range of knowledge, and have appropriate complexity and reasonable distribution. Secondly, the test bank must be easy to operate so that the teachers can use it and maintain it intuitively. Plus, the test bank should also be open and extendable so as to allow the teachers to add in

new test questions and adjust the principle of automatic test paper. According to the above principles, the author designed and developed the “engineering graphics test bank”.

2. THE MAIN FUNCTIONS AND FEATURES OF THE TEST BANK

The main functions are shown in figure 1.

The main characteristics of the test bank include the following:

(1) The management system of the test bank is rationally designed with complete functions, reliable performance, high fault tolerance, friendly interface ,easy operation as well as good open-ness and extendibility. So it is easy to be used and maintained by teachers.

(2) The test bank has complete and accurate property settings of test questions, including: Question type, Complexity, Tested knowledge, Test source, Similarity factor, Estimated duration and Allotted score, etc. Among these property settings, “test source” helps users decide whether to adopt test questions from text book; “similarity factor” helps to avoid similar questions appearing on a same test paper. The test content is complete and standard, including test questions, answers and marking criterion.

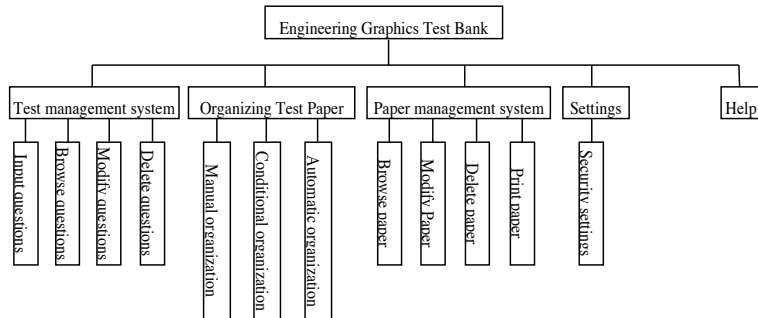


Figure 1: Engineering graphing test bank system

(3) Flexible and diverse ways of paper organization are available, including manual organization, conditional organization and automatic organization. Automatic Organization has high practical value for its advantages such as complete design, intuitional interface and easy operation.

(4) The system can save the built-up papers and the users can choose to preview, modify, delete or print the saved papers.

(5) The system has a intelligent function of setting type and can automatically arrange the pages so as to produce papers with complete content (including title, score distribution list box, test requirements, and the respective number broad theme questions and narrow theme questions). The overall arrangement is rational and needless of too much human intervene.

(6) The system also has a function of test questions statistic, confidentiality setting function and exhaustive help.

3. STUDY OF ORGANIZING TEST PAPER

The management system of auto-organizing test paper provides users with functions of viewing and selecting test questions and organizing papers in certain ways. Flexible and diverse paper organization methods are available, including manual organization, conditional organization and automatic

organization.

When the paper organization is entered, what emerges before users is an interface where they can type in paper information such as school name, course title, major, applicable class hour, and test time and so on, all of which appear in the top zone of test paper. And when a new value is entered in the “school name” option box, users can choose whether to “set it as default value” or not. If they do, the “school name” box will show the new default value when they enters the paper organization system next time.

3.1 Manual Paper Organization

Manual paper organization is an organization in which the users organize papers by viewing the test bank through conditional search and selecting test questions that meet their requirements one by one. Its advantage is that the users have complete control of the properties of every test question and know them fairly well, while the disadvantage is higher time consumption and lower efficiency.

3.2 Conditional Paper Organization

Conditional paper organization refers to the system automatically selects a specified number of test questions that meet the requirements set by the users and organize test papers.

The difference between manual paper organization and conditional paper organization

is that the system automatically searches for test questions based on the requirement set by the users, from which it then randomly selects a specified quantity of test. After the test questions are selected, the users can view and modify them, for instance, they can manually delete a certain question if they found it unsuitable, or they can add a new test questions to the paper. And whenever they are satisfied with the test questions, the system will organize the paper and produces the answers. The advantage of conditional paper organize is that the teacher can have fairly good control of the properties of test questions and ensure the objectivity of test paper at the same time, allowing for convenience and quality at the

same time.

3.3 Automatic Paper organization

Automatic paper organization is based on the pre-set project to produce complete test papers and features the advantage in convenience and swiftness. In order to make automatic paper organization more rational and practicable, it is necessary to ensure appropriate distribution of test knowledge, proper difficulty, proper question types and reasonable score settings plus convenience to modify. And the test bank designed is just based on these requirements. The structure of this test bank is shown in figure 2.

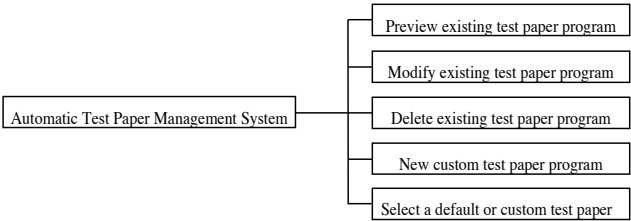


Figure 2: Automatic Test Paper Management Module

This system provides four course options including descriptive geometry course, engineering graphing course, mechanical graphing course and modern engineering graphing, and each has three difficulty factors from easy, intermediate to difficult, thus there are altogether 12 automatic paper Organization default options. These default options are based on abundant teaching and testing experience, the characteristic of each course, the requirements of syllabus and the key point of each course. And for the purpose of making automatic paper organize more rational, the default values include factors such as the distribution of test questions, the distribution of knowledge in each chapter, the difficulty factor and allotted score of each question, etc.

In the management module of automatic paper organize system, users can preview the default solutions or user-defined solutions, and

have access right to modify or delete user-defined options, but only system administrator have access to deleting or modifying the default options.

Users can also create new options on the basis of default options according to the differences in major, course hours and teaching objects. The interface of this function is shown in figure 3.

The option boxes in this interface is based on the most usual question type(simply “QT” for short in this article) of the related course, the columns including tested knowledge(simply “TK” for short in this article) of each question type, while the rows including the difficulty factor, quantity and average scores corresponding the knowledge involved in question type in the columns. If the users choose the existing foundation solutions, then the values appear in their relevant positions; if

users choose a blank solution, then all the option boxes are empty.

The option boxes allow the users to modify or input in the knowledge responding to each question types, quantity of test questions and average score allotment. As users type in these values, the “The paper’s setting list” on the screen shows the quantity of the selected question type, allotted score and surplus scores in real time so that users can have full knowledge about test paper organization.

The user-defined quantity of test questions cannot exceed the existing quantity of related question types; otherwise there will pop up an error display. To prevent this from happening,

the system includes a function of “real time test question counting”, so that users need only to click the option boxes of relevant test questions with a short pause to have the total test questions number of the related question type, as shown in Figure 3. This function helps users define valid quantity of test questions of each question type.

In all of the three paper organization methods above, users can save the files for later check, modification, type setting and printing. Users can also enter the interface of “View Paper” to modify, delete, type set and print the previously organized papers.

Tested Knowledge	Question Quantity			Score	Tested Knowledge	Question Quantity			Score
	Easy	Medium	Difficult			Easy	Medium	Difficult	
TK1	0	0	0	0	TK10	0	0	1	12
TK2	0	0	1	0	TK11	0	0	0	0
TK3	0	0	0	0	TK12	0	0	0	0
TK4	0	0	0	0	TK13	0	0	0	0
TK5	0	1	0	7	TK14	0	0	0	0
TK6	0	0	0	0	TK15	0	0	0	0
TK7	1	1	1	7	TK16	0	0	0	0
TK8	0	1	0	0					
TK9									

The paper's setting list

I. QT1
4 questions, 8 points

J. QT3
4 questions, 8 points

H. QT4
2 questions, 12 points

H. QT5
2 questions, 12 points

V. QT6
1 questions, 5 points

H. QT7
7 questions, 12 points

A total of questions: 23
The surplus score: 0

Return to the program

Save

Return to the main interface

Exit

Figure 3: New custom test paper program

4. CONCLUSIONS

The test bank system is designed properly with abundant test questions that have broad representativeness and good optionality as well as diversified and flexible paper organization pattern. In particular, the automatic paper organization module has complete functions, intuitional interface, easy operation and convenience to modify. This software has been

put into practice successfully.

ACKNOWLEDGMENTS

I would like to express my gratitude to Jin Dang and my colleagues for giving me help during the thesis writing process, also be very grateful to Hans-Peter Schröcker and other staff working for 16th International Conference on Geometry and Graphics.

REFERENCES

- [1] Kong Xianshu, Zhang Li, Zhong Jun. Descriptive Geometry and Mechanical Drawing Test Bank System [A]. Graphics Education Research[C]. Beijing: Mechanical Industry Press, 1998.194-200. (In Chinese)
- [2] Luo Jun, Jiao Yong-he and Yang Wei. Discussion about Building General Test Bank Management System [J]. JOURNAL OF ENGINEERING GRAPHICS, 2002, 23(2): 157-161. (In Chinese)

ABOUT THE AUTHORS

- 1. Lijie Guan is a professor in the College of Mechanical Science and Engineering in Northeast Petroleum University, Daqing, China. Her research interests are graphics education, CG and 3D Modeling. She can be reached by e-mail: gljws@163.com, by Fax: +86(0459)6503740, by phone: +86(0459)6503740, by address: College of Mechanical Science and Engineering, Northeast Petroleum University, Daqing, Heilongjiang Province, 163318, P.R. China.
- 2. Huilai Liu is a master student in the College of Mechanical Science and Engineering in Northeast Petroleum University, Daqing, China. His research interests are CG, 3D Modeling and Mechanical design. He can be reached by e-mail: 546985030@qq.com, by Fax: +86(0459)6503740, by phone: +8615845987580 or through the postal address: College of Mechanical Science and Engineering, Northeast Petroleum University, Daqing, Heilongjiang Province, 163318, P.R. China.
- 3. Chunhua Wang is a professor in the College of Mechanical Science and Engineering in Northeast Petroleum University, Daqing, China. Her research interests are graphics education, CG, 3D Modeling. She can be reached by e-mail: ztwch@126.com, by Fax: +86(0459)6503740, by phone: +8613936911258

or through the postal address: College of Mechanical Science and Engineering, Northeast Petroleum University, Daqing, Heilongjiang Province, 163318, P.R. China.

- 4. Feng GUO is a professor in the College of Mechanical Science and Engineering in Northeast Petroleum University, Daqing, China. Her research interests are graphics education, CG, 3D Modeling. She can be reached by e-mail: gfgrace@126.com, by Fax: +86(0459)6503740, by phone: +8613394599886 or through the postal address: College of Mechanical Science and Engineering, Northeast Petroleum University, Daqing, Heilongjiang Province, 163318, P.R. China.

RIGID QUADRILATERAL FOLDING STRATEGIES FOR SURFACE DESIGN IN ARCHITECTURE

Albert WILTSCHE and Milena STAVRIC
Graz University of Technology, Austria

ABSTRACT: In the field of origami it has been investigated how to design different spatial shapes only from one sheet of paper using folding techniques. In folding techniques single planar material is folded without stretching, tearing or cutting. This theoretical definition exists only in geometrical and mathematical context where origami forms are regarded as ideal zero-thickness surfaces. If we consider material thickness and stiffness (e.g. cardboard, plastic sheet or sheet metal) for large scale space structures, stretching and tearing inevitably appear in the construction. If we use building material with a certain thickness in real architectural size, corrugated structures cannot be folded but the shape of individual parts can follow a folding pattern. The aim of our research is to transfer folding principles to buildable structures made of standard materials. Our approach is not to make moveable and foldable structures but to use geometrical rules and advantages of folded surface structures to create large span load bearing structures which follow a certain geometric shape.

Keywords: Rigid Folding, Origami, Single Curved Surfaces.

1. INTRODUCTION

In our previous research we investigated origami folding patterns like Yoshimura Pattern, Diagonal pattern and Muira Ori pattern in order to find geometrical properties and rules that can be used in architectural context ([9]). We developed parametrical models with zero-thickness that allowed us to analyse different patterns and to make scale models in order to understand limitations and the potential of folding structures. Folding structures enables a high degree of constructional efficiency by reducing the weight of the structural elements simultaneously.

Our proposed geometrical surfaces are single curved surfaces which are discretized into quadrilateral planar elements that follow up folding principles. Starting with quad mesh folding pattern, which follow cylinders, we will transfer the construction to cones of revolution and developable helical surfaces. With some adaptations we can also transfer the results to forms which follow surfaces of revolution and

sweeping surfaces.

In our approach we use geometrical rules in the form-finding phase. The out coming results are transferred to the singular elements which are cut with miter and afterwards connected accurately in a structure. We investigate the possibility of offsetting and we analyze different parameters that influence offset and meets the requirements in case where four parts meet in one node.

Our approach is not making classical kinetic foldable structures but to use geometrical principles and static advantages of origami in order to develop interactive methods for designing rigid structures with a load-bearing effect for different geometrical structures with quadrilateral patterns.

The theoretical basis of our work is funded in various research fields which cover geometrical and mathematical properties of classical origami [1],[2],[3],[7],[10] and folding structures and principles from the mathematical point of view [4],[5],[6],[13].

2. ARCHITECTURAL PRACTICE

Folding structures can be found in many fields, such as industrial design, fashion, interior design, architecture, textile industry and jewelry. They can be made out of different types of material, such as paper, textiles, cardboard, wood or metal. In architectural terminology, the term folding structure stands for structures consisting of plane polygonal elements.

There already exist some very interesting architectural examples which demonstrate the potential of structures developed from folding structures like the Folded-Plate Hut in Osaka by Ryuichi Ashizawa Architects, Yokohama International Port Terminal in Tokyo by Foreign Office Architects, United States Air Force Academy – Cadet Chapel by Leo A. Daly, Inc. and Henningson and the Chapele St. Loup in Switzerland by Danilo Mondada.

2.1 Simplifying the system

In general a folding pattern for three-dimensional folding consists of triangles or/and quadrangles. Patterns with triangles produce a great variety of folding shapes (figure 1) but implicate two big problems. A knot of a triangle folding pattern is usually the centre of six parts meeting there. This causes much trouble for architectural production which is usually processed in big scale. The second disadvantage is the geometrical control of a triangle folding system as there are too many degrees of freedom.

In this sense we concentrate on folding patterns and systems consisting of quadrangles and knots where four of them are meeting.



Figure 1: Triangle folding form (student's project).

2.2 Material efficiency

From the architectural point of view there are some important facts which make the use of folding structures interesting:

Foldable material can be used to cover a whole construction in a certain range with only one piece (Figure 2 left and middle). For instance a simple demountable frame construction can be easily covered by one piece of canvas like a tent. As a second application one piece of plastic layer can be used to laminate material to make it resistant over atmospheric exposure.

A spatial structure developed by folding which is made of a non-foldable material (i.e. timber panel) leads to a "natural" optimal nesting of the parts for their production (Figure 2 right).

Folding structures can be used to improve the static behavior of loads.

If the geometry of folding structures is clear, it is easy to program a parametric model which leads to a comfortable way of design.

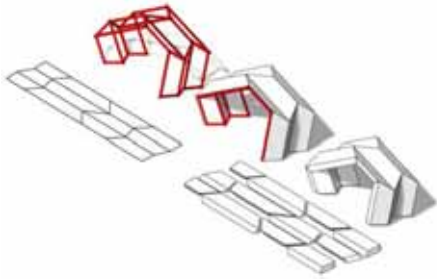


Figure 2: A frame construction can be covered only by one piece of material (left and center). Constructions with thicker material lead to a natural nesting.

2.3 Offset

There is one big problem for the architectural practice which is based on geometry and it is affected by material thickness. The geometrical translation of this problem is the term offset. If one takes four or more planes and puts them together in one point then the offset of the planes (= thickness) in general does not meet in a point (figure 3). Besides aesthetic problems there is also a static imbalance in such a knot which can cause difficulties in the loading capacity. To avoid these troubles there is a simple geometric solution. If all the involved planes are tangent to a common sphere or cone of revolution, the correct offset is guaranteed because the miter planes between two neighboring planes all meet in one line ([9], [12]). This line connects the knot and the centre of the sphere or rather is identical with the axis of the cone. A much more practical way of computing and programming is given in [11]. If four planes meet in one knot the offset is correct iff the opposite angles in the knot add up to the same sum.



Figure 3: The offset problem. The initial developable folding shape (left) gets an offset (right) which is not practicable.

3. FOLDING TYPES

We will distinguish between two folding types – linear and curved folding. Figure 4 left shows a linear folding structure that is developed by folding a plane material along (straight) edges. This type consists usually of triangles and quadrangles. A curved folding structure is built by folding a plane material along curves (figure 4 right). It consists of single curved surfaces, such as cylinders, cones and general developable surfaces. For the following we will only investigate the linear folding type.

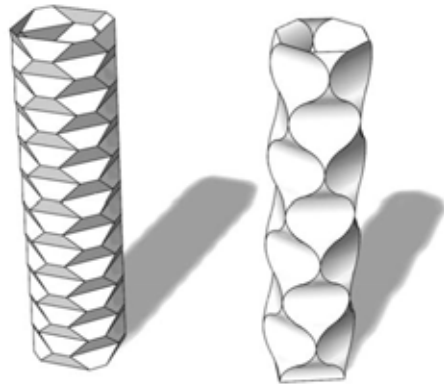


Figure 4: Linear and curved folding. The example on the left side is composed of trapezoids, the one on the right side of cylinders.

Since some materials are very flexible, a folding structure can be further bended or twisted (figure 5). In this case the plane ele-

ments are no more flat but curved and it is almost impossible to predict the geometrical form. We will only investigate linear folding structures whose plane elements stay plane throughout the whole folding act. This term is called rigid folding.

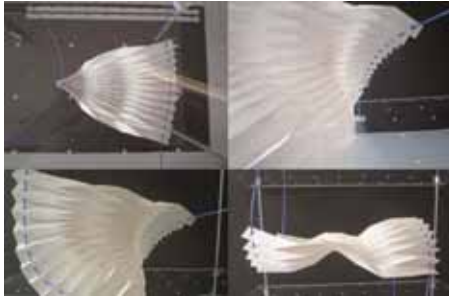


Figure 5: A folding shape which is forced in different positions. The initial flat quadrangles are no more flat, so this is no rigid folding.

4. FOLDING TECHNIQUES

4.1 Types

If a sheet of paper is folded along parallel edges the folding is called parallel folding. This form appears in architecture context very often although the intention of the original design might not have been a folding shape. Figure 6 shows a form similar to a shed roof and a folded form which approximates a cylinder. The left form changes always the orientation of the parallel folds. This folds are called mountain and valley folds. The right form is folded continuously in the same orientation.

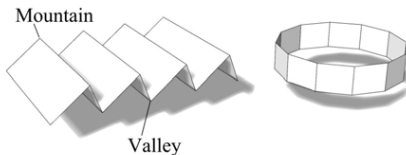


Figure 6: Parallel folding, mountain and valley folds.

A second type which is much more interesting in the case of geometry and designing is called reverse folding. The initial form is a parallel folded shape which is split by an arbitrary plane. Then one of the parts is mirrored on the splitting plane (figure 7 upper row). Figure 7 bottom shows an arbitrary parallel mountain valley folding designed by the reverse method.

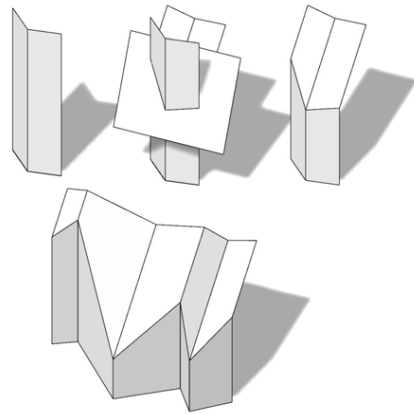


Figure 7: Reverse folding

4.2 Reverse folding

There exists an additional interesting fact concerning reverse folding. Given a flat pattern consisting of parallel strips and a polygon e with vertices on the strip edges as shown in figure 11 left. If the pattern is folded like in figure 11 right, then the folded polygon e lies always in a plane ϵ .

Figure 8 left shows a plane α and two different intersecting lines e and g on it. If α is folded along the edge e the plane is split up into two planes α_1 and α_2 and the line g is divided into g_1 , g_2 and determines a plane ϵ as long as the angle of g_1 , g_2 is not 0° or 180° . The planes α_1 and α_2 have two symmetry planes and we are using the plane σ which is shown in the figure. If we intersect σ and ϵ we get the unique line n which is a symmetry line of g_1 and g_2 . A rotation around n about 180° transforms g_1 into g_2 but also α_1 into α_2 .

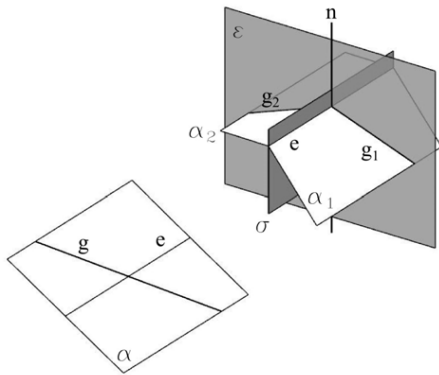


Figure 8: Folding a plane α around a line e . The outcome is two planes α_1, α_2 and two lines g_1, g_2 .

The flat folding pattern of figure 9 shows the line g already split up in the two parts g_1, g_2 and two folding lines e and f . The initial flat position may be covered by two coincident planes α and β . If we fold this pattern around e and f we get the form shown in figure 10 left. The flexed parts g_1, g_2 are now simultaneously parts of the folded planes α and β .

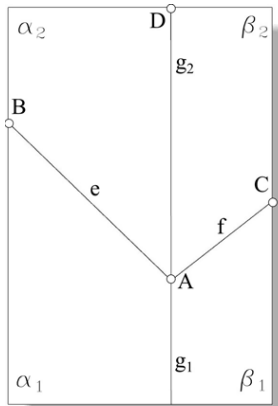


Figure 9: Flat pattern with folding edges e, f, g_1 and g_2 .

As said before there exists a unique sym-

metry line n of g_1 and g_2 which now must be orthogonal both to e and f . On the other hand n must be also the intersection line of the appropriate symmetry planes of α_1, α_2 and β_1, β_2 respectively. This means that the direction of n is determined.

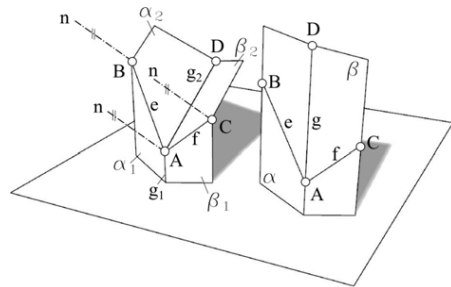


Figure 10: Two Folding shapes of the pattern of figure 9.

This means that a folded version of connected parallel stripes with edges in the direction of g and a polygon e with vertices on the edges provides a folded polygon e which lies always in a plane ϵ orthogonal to a direction n . The direction n can be found intersecting appropriate symmetry planes.

Figure 10 right shows another folding position where g, α and β are not split.

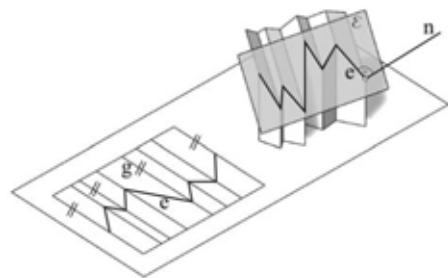


Figure 11: Every polyline p with vertices on parallel lines g stays in a plane ϵ for any arbitrary folding position.

5. SWEEP FOLDING

In order the get a great variety of shapes the

idea of reverse folding can be extended in a way that is similar to the idea of sweeping. Sweeping is a method of generating surfaces by a profile and a rail curve. It is a well-known feature in CAD modeling packages. The input for our “sweep-folding” is a “profile-polygon” and a “rail-polygon” and the profile is moved along the rail. In general the two curves are plane and the planes are orthogonal but not necessarily. This act would generate a typical sweeping surface, which cannot be built by folding. In order to get a folding shape, a reverse folding must be performed in every vertex of the rail-polygon. The corresponding mirror plane σ_i in a vertex V_i is one of the two symmetry planes of the adjacent rail edges $r_{i-1,i}$ and $r_{i,i+1}$. Figure 12 shows this designing principle. The profile polygon p lies in a plane orthogonal to related rail edges.

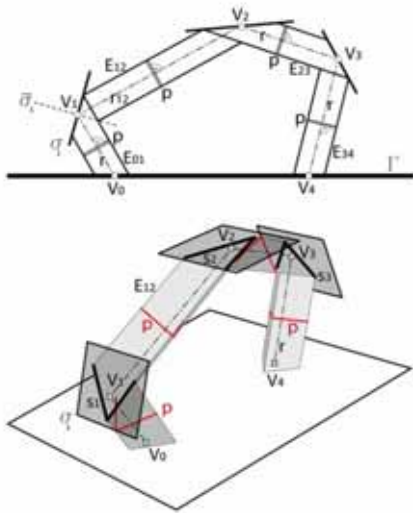


Figure 12: Sweep-folding: A profile-polygon p is moved along a rail-polygon r . At the vertices V_1 , V_2 and V_3 the principle of reverse folding is performed.

Between two rail vertices V_i and V_{i+1} the profile p and the rail edge $r_{i,i+1}$ generate an extru-

sion surface $E_{i,i+1}$. At the vertex V_i the surface $E_{i,i+1}$ is cut by the appropriate symmetry plane σ_i . The second part of $E_{i,i+1}$ is then mirrored by σ_i to get the next part $E_{i+1,i+2}$. At the start and end vertex of the rail r the folding shape can be cut arbitrarily. Figure 12 bottom shows a general view of the situation. There it can be seen that the intersection polygons s_i of the extrusion surfaces $E_{i-1,i}$ with the symmetry planes σ_i can be interpreted as projections of the profile curve onto it in direction of $r_{i-1,i}$. This fact leads also to a simple possibility to construct the whole shape as a loft surface between the polygons s_i .

If the profile is extended or copied several times orthogonal to the plane of the rail a shape approximating a cylindrical surface can be generated. Figure 13 shows such a form following a spiral cylinder.

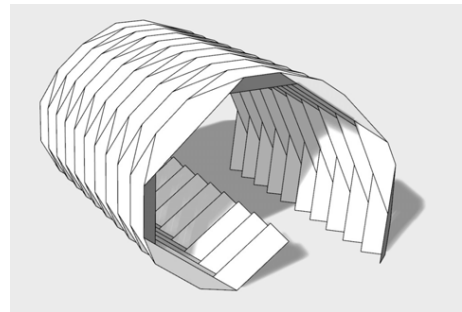


Figure 13: A sweep folding shape following a cylinder with a spiral as profile curve.

Further variations of this generation principle can be made by rotating the profile curve in its plane. Another variation is to use three-dimensional curves. This makes only sense for the rail. Figure 14 shows such a 3D-rail, a plane profile curve and the associated foldable form. In each vertex of the rail curve the sweep folding form is cut by the appropriate symmetry plane belonging to the neighboring rail edges.

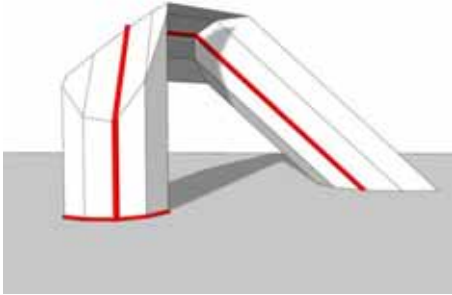


Figure 14: 3D-rail.

6. CONICAL FOLDING SHAPES

In order to design folding shapes which follow a cone we will show a construction principle deduced from a work of H. Stachel ([8]). Stachel shows that the folding of an ornamental pattern consisting of congruent and convex arbitrary quadrangles generates a form which follows a cylinder of revolution, or rather all pattern vertices lie on this cylinder. The quadrangles can be transformed into each other by screw motions.

If we now take trapezoids whose vertices lie on a cone and transform them by spiral motions instead of a screw motion we can design a form which follows a cone. Figure 15 left shows different discrete positions of an initial trapezoid generated by a spiral motion.

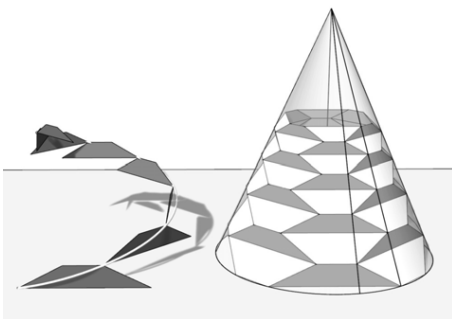


Figure 15: Conical folding shape generated by a spiral motion of trapezoids.

Figure 15 right shows the folding form following a cone. There are two different shapes

of trapezoids, whose vertices lie on the cone.

7. SURFACES OF REVOLUTION

Surfaces of revolution can be approximated by cones of revolution with the same axis. We can replace the cones by the folding forms shown in the previous chapter (figure 16). But unlike the conical case a surface of revolution cannot be folded by a connected piece of material. The flat folding pattern consists of curved stripes each forming a different cone for the initial surface.

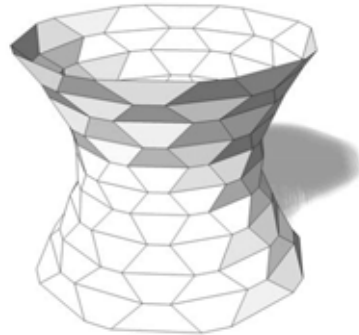


Figure 16: A folding form approximating a surface of revolution.

8. HELICAL SURFACE

Similar to the conical case folding patterns for developable helical surfaces can be created. In this case the unrolled folding pattern is again one connected piece of material. Figure 17 shows such a form consisting of trapezoids which can be transformed into each other by helical motions.

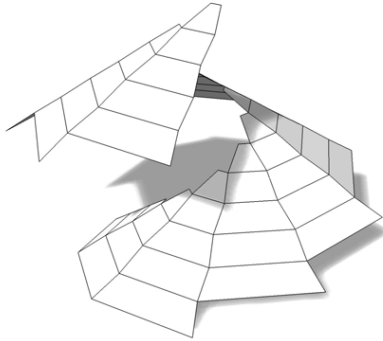


Figure 17: A folding form following a helical developable surface.

9. CONCLUSIONS

In this paper we discussed folding forms which follow different geometric shapes. For the folding we use only patterns consisting of quadrilaterals which approximate cylindrical and conical shapes but also discretized sweeping shapes, surfaces of revolution and developable helical surfaces. We showed the principle of reverse folding and the folding of an arbitrary polygon with vertices on the edges of parallel stripes which leads also to reverse folding.

All the shown examples should enable a creative architectural approach to the matter of folding and enlighten the jungle of endless folding shapes by use of “straight geometry”.

REFERENCES

- [1] S.M. Belcastro and T.C. Hull. A mathematical model for non-flat origami, in *Origami 3, Proceedings of the 3rd International Meeting of Origami Mathematics, Science and Education*: 39–51, 2002.
- [2] S.M. Belcastro and T.C. Hull. Modelling the folding of paper into three dimensions using affine transformations, *Linear Algebra and its Application*, 348: 273–282, 2002.
- [3] R. Geretschlaeger. *Geometric Origami*, Arbelos, UK, 2008.
- [4] T. Kawasaki. On the relation between mountain-creases and valley creases of a flat origami (abridged English translation), in: *Proceedings of the First International Meeting of Origami Science and Technology*, 229–237. Ferrara, 1998.
- [5] J. Mitani. A Design method for 3d origami based on rotational sweep. *Computer-aided Design and Application*, 6 (1): 69–79, 2009.
- [6] S. Miyazaki, T. Yasuda, S. Yokoi and J. Toriwaki. An origami playing simulator in the virtual space, *The Journal of Visualization and Computer Animation*, 7(1): 25–42, 1996.
- [7] T. Nojima. Modelling of folding patterns in flat membranes and cylinders by origami, *JSME International Journal Series C*, 45(1): 364–370, 2002
- [8] H. Stachel. A Flexible Planar Tessellation with a Flexion Tiling a Cylinder of Revolution. *Journal of Geometry and Graphics*, 16 (2): 153–170, 2012.
- [9] M. Stavric and A. Wilsche. Quadrilateral Patterns for Rigid Folding Structures. *International Journal of Architectural computing*, to appear.
- [10] T. Tachi. Generalization of Rigid-Foldable Quadrilateral-Mesh Origami, *Journal of the International Association for Shell and Spatial Structures*, 50(3): 173–179, 2009
- [11] W. Wang, J. Wallner and Y. Liu. An Angle Criterion for Conical Mesh Vertices. *Journal of Geometry and Graphics*, 11 (2): 199–208, 2007.
- [12] A. Wilsche, H. Schimek and M. Stavric. Geometric Aspects in Producing Non-Standard Architecture with Standard Tools. *Journal of Geometry and Graphics*, 12 (2): 205–214, 2008.
- [13] H. Zimmer, M. Campen, D. Bommers and L. Kobbelt. Rationalization of trian-

gle-based point-folding structures, *Computer Graphics Forum*, 31(2): 611–620, 2012.

ABOUT THE AUTHORS

1. Milena Stavric studied architecture. Her focus is on applied geometry in architecture.
2. Albert Wiltche is trained in geometry and mathematics. His focus is on geometry in architecture.

ON THE ROLE OF CIRCULAR SECTIONS OF QUADRIC SURFACES

The elaboration of the topic by two creative geometric student's tasks

**Aleksandar ČUČAKOVIĆ, Magdalena DRAGOVIĆ, Luka LAZAREVIĆ and
Đorđe NEDELJKOVIĆ**

The University of Belgrade, Serbia

ABSTRACT: As a part of new advanced course curriculum of descriptive geometry, in the Faculty of Civil engineering, creative 3D modeling of specific structures is included. These structures, inspired by contemporary architecture achievements, relay on the basic theory of quadric surfaces (also elaborated in the course - Mathematics 2). Throughout two given 3D modeling tasks, students elaborate two separate subtopics: discretization of the surface elliptic hyperboloid of one sheet (ELHY) and spiral surfaces, both connected by key elements – circular sections of quadrics. The subtopic's methodology designed for student's tasks is guided by 3D structures modeling process through applications (the first topic-Auto Lisp app./ the second topic-Visual Basic app.) implemented in engineering software *Auto CAD* and final advanced 3D modeling operations.

Keywords: students geometry task, circular sections of quadric surfaces, polygonal mesh structure, spiral surfaces.

1. INTRODUCTION

The educational necessity of any technical i.e. engineering faculty nowadays concerns 3D modeling of various structures, depending on the field of interest. Various investigations on this subject [2, 9, 10] were carried out. The 3D modeling has an important prior – designing role before final prototyping [2]. The 3D modeling is significantly present when considered a variety of available parametric modelers: *Rinoceros 3D*, *3Dmax*, *Solid Works*, *Catia* and many others. The inconvenience of such software solutions, when issue of Descriptive Geometry is of the matter [10], is that real understanding, i.e. recognizing of the geometric procedures is reduced, almost neglected, in order to obtain the final "product".

The 2nd year curriculum of the basic studies on the Civil Engineering Faculty enables the knowledge earning and linking, concerning several areas: descriptive geometry, algebraic geometry, computer drawings (graphics) and computer programming. Regarding several

reasons, especially to get closer the theory and engineering practice [8], the advanced descriptive geometry course, in the Faculty of Civil Engineering in Belgrade, implemented the topics which elaborate challenging 3D modeling tasks. This initiative should be encouraging for the students in their individual modeling design of further scholar projects (buildings, bridges, or other civil engineering structures). Some extraordinary examples from engineering practice make this topic worth of attention (Fig.1).



Fig.1 Spiral bridge, Singapore [12]

In the course, the modeler is *Auto CAD* software, commonly used in engineering practice, in the role of teaching instrument, as well as a very precise and efficient tool. Since this is the first students' experience with modeling,

significant technical support is provided by two applications written in Auto-Lisp and Visual Basic programming languages.

In order to be actual in design, the chosen topics concern the quadric surfaces (simultaneously elaborated in students' course Mathematics 2) and as an extension – the spiral surfaces. The prior descriptive geometry approach emphasizes the idea of better spatial understanding of the quadric surfaces' characteristics along with spiral structures, commonly present in the form of algebraic (the quadrics) and parametric (the spiral surfaces) equations. Both are supported with computer applications aimed for modeling process where the final results are creative 3D structures

2. TWO GEOMETRIC TOPICS

The first creative theme is *discretization of ELHY surface*, i.e. generation of polygonal mesh structure, offering three solutions: from simplest – triangular, quadrilateral, up to the concave – hexagonal "ornaments" [9]. This topic has a rather simple explanation and graphical presentation (wire-frame and surface models) of geometrical procedures, i.e. constructions, in opposite to some complex analytical geometry solutions aimed for the parametric modeling [5, 6].

The second topic concerns various solutions of *spiral surfaces* deriving from geometric origins - quadric surfaces: cone, sphere, ellipsoid, elliptic hyperboloid and paraboloid. The regular polygons, inscribed in circular sections of the quadrics in spiral "motion" obtain the spiral surface/structure model. Two types of structures are offered, depending on the line of centers of inscribed polygons.

Both topics are elaborated through student's tasks, where the final result is 3D Auto CAD model (prototype) of the structure with specified (practical) aim.

2.1 Task 1 - ELHY surface discretization

The method of generating the ELHY's wire-frame model relies on settings of two arbitrary circular sections (center points - C and C_1) in two parallel horizontal planes and one

generatrix AB (Fig.2). Detail elaboration of the procedure is given in previous author's research, with accompanying basic CAD application [3, 4]. In accordance to intention of enabling efficient tool for discretization of ELHY shaped structure, the extensions of basic application were created for the purposes of the student's practicing. The basic application generates ELHY's net (two series) of generatrices and stores entity handles in the external file, that is used by the application extension. The extension offer three different discretization solutions-three polygonal meshes: triangular, quadrilateral and hexagonal, with planar elements.

The *first option*, the simplest one, concerns triangulated ELHY's mesh obtained by circular sections and the net of generatrices (Fig. 3) The ELHY surface is divided on horizontal "rings" by circular sections. The intersection points of generatrices – nodes are connected into planar triangles, varying in dimensions. All the triangle's edges contained in separate "ring" are equal. Their dimensions depend on the number of generatrices and the radius of the circular section as well.

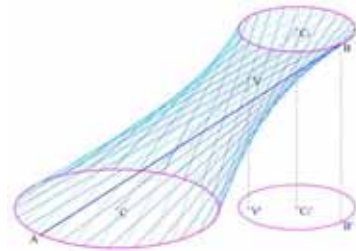


Figure 2: Wire-frame model of ELHY

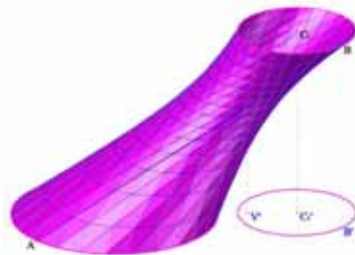


Figure 3: Triangular ELHY's mesh

The modeling task for mesh prototype is supported by basic Auto Lisp application - wireframe model of generatrices and circular sections accompanied with surface model.

The *second option* relies on the wireframe model also. In this case circular sections are obtained by the coplanar midpoints of neighbor nodes on each generatrix. The ELHY surface is divided on horizontal strips by circular sections. In each stripe trapezoids and parallelograms alternate each other. In fact, a tangential plane of each node is bounded by trapezoid; between trapezoids a sequence of variable parallelograms lays. (Figs. 4, 4a)

The application for generation of quadrilateral ELHY's mesh uses nodes from triangular mesh to calculate circular sections which contain sides of trapezoids and parallelograms. When all nodes for mesh generation are calculated, application rearranges them in specific order to provide automation of generation process. Application offers the possibility to choose different layer (color) for parallelograms, so they can be easily discerned in the mesh.

The *third option* offers hexagonal (concave) solution of ELHY's mesh. The specific disposition of nodes is set, obtaining one structural cell, which is multiplied. One central node is surrounded by 6 neighbor nodes (convex hexagon disposition) incorporated in the predefined basic ELHY's net of generatrices. Here, 7 tangential planes are bounded by 7 concave hexagons. (Figs. 5, 5a)

Each hexagon stretches through two horizontal ELHY's "rings". Intersection lines – the edges of hexagon are imagery diagonals stretching through three "rings" while connecting basic nodes. Besides, the geometric regularity of alternation of tangent line and secant, passing through the coplanar nodes (contained in separate circular sections) is utilized for practical Auto Lisp application.

The application for generation of hexagonal ELHY's mesh "walks" through each circular section and, using the above mentioned principles, calculates nodes for the mesh generation.

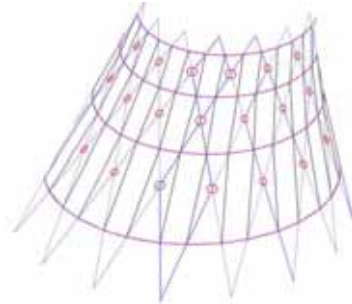


Fig. 4 Detail of quadrilateral ELHY's mesh

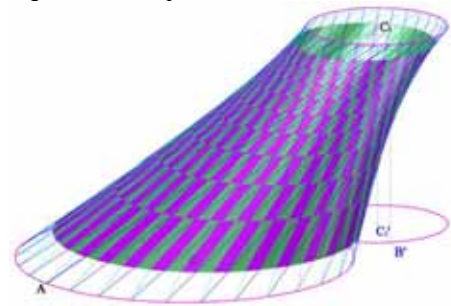


Figure 4a: Quadrilateral ELHY's mesh

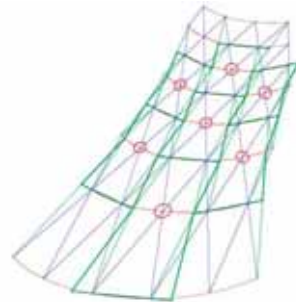


Figure 5: Detail of hexagonal ELHY's mesh

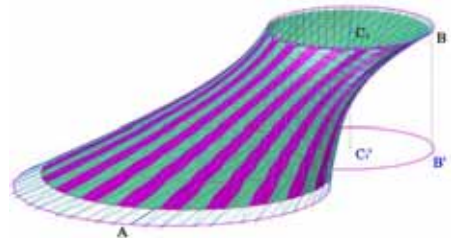


Figure 5a: Hexagonal ELHY's mesh

This application also rearranges nodes in specific order to provide automation of mesh generation process.

2.2 Task 2 - Spiral structure

Spiral surfaces of the specific type, with geometric origins: cone, sphere, ellipsoid, elliptic hyperboloid and paraboloid are chosen "set" for the modeling process in Auto CAD software. Each surface has two series of circular sections, in this case used as a pattern for inscribed regular polygons – parallel sections of the final surface's model.

The *first solution* for the spiral surface wireframe model considered the option that the line of centers for rotated polygons coincides with the line of centers of the origin surface's circular sections. The designer's choice of parameters is only relevant for the resulting spiral surface model. (Fig. 6)

The *second solution* varied the line of centers for rotated polygons. Geometric relation of the first and the last position of the basic polygon, i.e. geometric concept, is predefined in application by adding a new center of rotation in the polygon's plane. Hence, depending on geometric concept, line of centers may be the "other" straight line (origin surface - cone), or a curve (origin surface-sphere (Fig. 7), ellipsoid, ELHY, etc.).

Technical supportive "tool" for geometric construction of the wireframe surface model is created for each type of the origin surface, i.e. each surface has its own VBA subroutine, which has to be uploaded as a separate application and run from the command line. The "user" – a student, after the task introduction and theoretical explanations, is guided through the settings of the initial parameters of the primary surface (e.g. values for three axes of the ellipsoid) with circular sections, and afterwards, the spiral surface (the polygon type and fill rotation angle, i.e. rotation "step" angle). The application routine is presented in Fig. 8.

Precondition for the application execution is VBA module installation. It has to be adequate

for Auto CAD software's version, as well as for operative system's architecture.

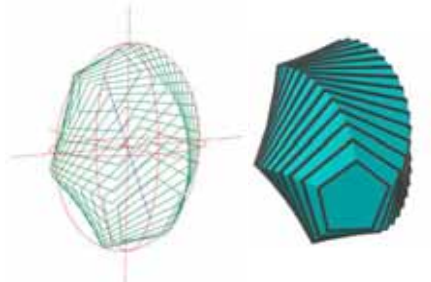


Fig.6 Wireframe and solid model of spiral structure E5 (ellipsoid/inscribed pentagons)



Fig. 7 Wireframe and solid model of spiral structure S4a (sphere/inscribed squares)

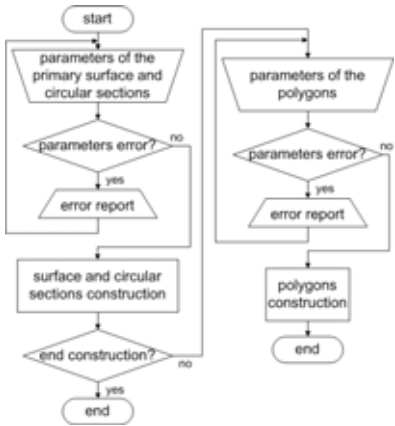


Figure 8: Application algorithm

No additional settings are needed for running the application inside drawing document. All constructive elements are set in separate layers. That makes easier further modeling procedure.

The procedure contains dialog boxes with textual descriptions of entry parameters and their ranges. If entered parameters out of the defined range, the user gets error response being returned to the entry (Fig. 9).

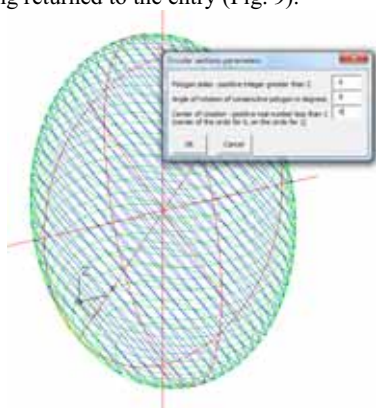


Figure 9: Dialog box for entering the spiral surface's parameters

Further modeling process includes utilization of advanced Auto CAD's modeling tools (e.g. "loft" command for twisted solid structure, or "sweep" command for the frames extrusion, Boolean operations, etc.)

3. DIDACTICAL PRINCIPLES

Geometrical principles presented in the first student task elaborate the interesting case of the ruled surface containing two series of straight lines – generatrices (rulings), where the solutions of discretization can be carried out by simple intersections of tangent planes, defined by nodes in various dispositions. Thus, geometric solution is "visible", i.e. obvious, and can be manually drawn directly in 3D computer surroundings, on the basis of the wire-frame surface (ELHY's) model. This is an important issue when the students' spatial abilities are of the matter. The final "product" in this task should be a complex structure obtained by combinations of shapes (identical or various, multiplied, cut, etc.), where spatial imagination and creation have an important role. Thereby, a student should present his modeling skills, by

proper utilization of Auto CAD's modeling tools.

Manual geometric construction of mesh structure unfortunately requires a lot of patience and time, and hence, for this purposes it is not convenient and practical. However, the possibility of making the "paper" surface model exists, regarding developable "rings", or "strips" on ELHY surface. Students' efforts in direction of exploring the possibilities in exercising with real (exact) models, recommended by P. H. Meier [11], should be stimulated and additionally estimated. In this case, the advanced task relies on the possibility of creating a new discretization pattern, i.e. engagement of new geometrical ornaments (regarding disposition of nodes) which is offered optionally for excellent students.

The didactical reasons, regarding understanding of descriptive geometric principles in solving of the second presented task, imply explanation of complete constructive and modeling procedure including the detail explanation of geometric procedure for obtaining the circular cross sections of 2nd order surfaces. Regarding practical reasons, such as minimization of manual drawing of circular sections, inscribed polygons, with numerous repetitions of necessary operations (rotation of the separate objects, UCS positioning, etc.) in Auto CAD's environment procedures, the design of appropriate VBA application is created.

The intention of such approach was to provide the continuity between spatial understanding of geometric procedures, analytical expressions of the geometric problem, computer programming, and finally, model designing and visualization. This concept affected the programming of the applications. The possibility to "stop" the "algorithm" is enabled, during separate parts explanations of the creative procedure. After each part of the construction the dialog box contains necessary details, i.e. coordinates and supportive tutorials.

The main reasons for the VBA programming language (Fig. 10) choice, in order to implement the application, are its possibility of run-


```

//axiomal msi plane
plane3D = Array(Var2D(Array(0, 0, 0), 0, 0), Var2D(Array(0, 0, 0), 0, 0),
Var2D(Array(0, 0, 0), 0, 0))
plane3D[1][0] = plane3D[1][0] + 1
plane3D[2][1] = plane3D[2][1] + 1

//draw ellipse
let objNewLayer = Layer("ellipse", axcenter, ellipseLayer)
currentAks = Var2D(Array(0,0,0), 0, 0)
currentCenter = Var2D(Array(0, 0, 0))

let i = 2 * Math.PI / num
for j = 1 to num - 1
  currentAks[i] = sqrt(1 - (1h - 1) * (int i / 50 - 2) * majorda * 2)
  currentCenter[i] = 1h - 2 * i * int
  let ellipse(i) = ThisDrawing.ModelSpace.AddEntity(currentCenter,
currentAks, ratio)
next

let objNewLayer(ellipse, center)

if majorda = 0 then Exit sub

mi = ratio * majorda * "semi-minor axis"

//construction of circle and ellipse
apoint = Var2D(Array(0, 0, 0))
apoint(1) = sqrt(mi * mi + majorda * majorda) / (mi * mi + 1)
apoint(2) = sqrt(mi * majorda * majorda * apoint(1) * apoint(1))
pum = apoint(1) / sqrt(apoint(1) * apoint(1) + apoint(2) * apoint(2))

//align
if pum <= 1 then
  alfa = atan(pum / sqrt(pum * pum + 1)) + 2 * atan(1)
  alignw = Tan(alfa)
  sub coordinate of the first circle on ellipsoid surface
  cpoint = Var2D(Array(0, 0, 0))
  cpoint(1) = sqrt(mi * mi + 1) * 0.5 / (sqrt(mi * majorda * mi + 1))
  cpoint(2) = -cpoint(1) * 0.5 * h / alignw * mi / md

else
  alfa = 0
  cpoint = Var2D(Array(0, 0, 0))
end if

```

4. CONCLUSIONS

ometric procedures for both topics is independent with respect to existing mathematically based software packages. Although the programming procedure requires analytical expressions, they strictly rely on the constructive geometric concept.

The expectations, within results of implementation of such challenging tasks in education process, are concerning of obtaining the continuity of various knowledge acquisitions towards practical applications in the future engineering practice.

Authors express their gratitude to the Ministry of Science and Technological Development of the Republic of Serbia, for supporting project No. TP 36008 entitled: *Development and application of scientific methods in design and building of high-economic structural systems by application of new technologies*, the part of which is the present study.

- [1] M. Contero, P. Company, J.L. Saorin and F. Naya. Learning support tools for developing spatial abilities in engineering design. *Mudd Design Workshop, V. Center for Design Education, Harvey Mudd College, Claremont, California, US, 19-21 May 2005.*
- [2] S. Converso. Mathematics for the design of variation: the "Nagashima!" lamp prototype. *Nexus Netw J*, 15 (3) 549-564, 2013.
- [3] A. Čučaković and M. Dragović. Geometric construction method of elliptic hyperboloid of one sheet using its circular sections. *Proc of 3rd Int Sci Conf moNGeometrija 2012 (Novi Sad, Serbia, June 21-24)* pp. 279-288. Faculty of Technical Sciences, Novi Sad.
- [4] M. Dragović and A. Čučaković. Architectural modeling - Elliptic hyperboloid of one sheet. *Spatium International Review*, to be published in July 2014.
- [5] V. N. Ivanov. Geometry and forming of the polyhedral box type surfaces based on the cyclic surfaces, *Structural mechanics of*

Engineering Constructions and Buildings,
No. 2, pp. 3-10, March 2012.

- [6] M. Manahl, H. Schimek, E. Ruffo, C. Dominguez and A. Wiltche, Ornamental discretisation of free-form surfaces, *Proc. of 17th Int. Conf. CAADRIA, 2012*. pp.347-356
- [7] F. Nassery. Auto CAD assisted teaching of descriptive geometry and engineering graphics. *20th Int. Conf. Geometria Grafika Komputer, 1-4 May, Wroclaw, Poland 2013*.
- [8] Y. Pei and S. Yan. The exploration and Practice of teaching method "Descriptive Geometry" and "Engineering Drawing" for art students. *19th Int. Conf. on Industrial Engineering and Engineering Management, 2013*. pp.1415-1421
- [9] M. Stavrić, U. Hirschberg and A. Wiltche, Spatializing planar ornaments – Towards esthetic control in segmenting building curved surfaces, *Proc. of 28th ECAADE conf.-Future Cities, Zurich, Switzerland, 15-18 Sept. 2010*. pp. 437-443
- [10] R. B. Štulić and J. T. Atanacković. Implementation of Computer Technologies in Descriptive Geometry teaching: surfaces of revolution, *Facta Universitatis, series: Architecture and Civil Engineering*, 2 (5) 379-385, 2003.
- [11] P. H. Maier, Spatial geometry and spatial ability - how to make solid geometry solid? In: Elmar Cohors-Fresenborg, K. Reiss, G. Toener & H.-G. Weigand (Eds.), *Selected Papers from the Annual Conference of Didactics of Mathematics 1996*, Osnabrueck, p. 63–75.
- [12] <http://www.prc-magazine.com/dna-inspire-d-bridge-wins-singapore-institute-architectural-award/>

ABOUT THE AUTHORS

1. Aleksandar Čučaković

PhD, architect, associated professor, Faculty of Civil Engineering, Department of Math, Phys, and Descriptive Geometry, University of Belgrade, Serbia. Fields of interest: descriptive geometry, perspective projections, projective geometry, computational geometry, geometry of surfaces, tessellation, geometric education, architecture, etc. E-mail: cucak@grf.bg.ac.rs

2. Magdalena Dragović

PhD, architect, assistant professor, Faculty of Civil Engineering, Department of Math, Phys, and Descriptive Geometry, University of Belgrade, Serbia. Fields of interest: descriptive geometry, projective geometry, computational geometry, geometry of surfaces, geometric education, architecture, etc.

E-mail: dim@grf.bg.ac.rs

3. Luka Lazarević

PhD candidate, civil engineer, Faculty of Civil Engineering – Department of roads, airports and railways, University of Belgrade, Serbia. Fields of interest: planning, design, superstructure, substructure and maintenance of railways, as well as application of CAD (Computer Aided Design) in design visualization.

E-mail: llazarevic@grf.bg.ac.rs

4. Đorđe Nedeljković

PhD candidate, civil engineer, Faculty of Civil Engineering – Department of Management, Building, Technologies and Informatics, University of Belgrade, Serbia. Fields of interest: automation in CAD, machine learning in civil engineering, etc. E-mail: ndjordje@grf.bg.ac.rs

THE ROLE OF HAND-DRAWING IN THE REPRESENTATION OF ARCHITECTURE AND LANDSCAPE

Maria Paola MARABOTTO
Polytechnic of Milan, Italy

ABSTRACT: The paper aims to highlight the role of drawing in the training of architects. Today, the Drawing is no longer seen as a mere instrument of representation, but it is a complete *medium* of knowledge, interpretation, reflection and awareness, not only of the environment surrounding the draftsman, but also his inner self. Thus, the importance and the role of design have established themselves in the field of education, broadening the scope of integrating with other disciplines, such as architectural design. Despite the progress of the new technologies, it is not possible to design without a basic understanding of drawing. In order for the idea of the architect to take shape and materialize, he must be able to transmit the image of a three-dimensional volume, from the mind to a two-dimensional paper sheet, or to the virtual space of a computer. He needs to work on that image to size, model, and create the architecture. The *Descriptive Geometry* provides the scientific basis of the representation's methods and therefore the readability of the drawing executed. With it, the student acquires the ability to represent space in two dimensions, to understand it, to verify it, and to size it. The geometric rules encoded transform the design into a universal language and lay the foundation for the various purposes of the drawing: design, surveying, knowledge, history, analysis, and communication.

Keywords: Still-life drawing, perception, knowledge, representation.

The paper aims to highlight the role of drawing in the training of architects. Today, the Drawing is no longer seen as a mere instrument of representation, but it is a complete *medium* of knowledge, interpretation, reflection and awareness, not only of the environment surrounding the draftsman, but also his inner self. Thus, the importance and the role of design have established themselves in the field of education, broadening the scope of integrating with other disciplines, such as architectural design. Despite the progress of the new technologies, it is not possible to design without a basic understanding of drawing. In order for the idea of the architect to take shape and materialize, he must be able to transmit the image of a three-dimensional volume, from the mind to a two-

dimensional paper sheet, or to the virtual space of a computer. He needs to work on that image to size, model, and create the architecture. The *Descriptive Geometry* provides the scientific basis of the representation's methods and therefore the readability of the drawing executed. With it, the student acquires the ability to represent space in two dimensions, to understand it, to verify it, and to size it. The geometric rules encoded transform the design into a universal language and lay the foundation for the various purposes of the drawing: design, surveying, knowledge, history, analysis, and communication.

A drawing executed by hand is an expression of representation. It is the foundation of the discipline because it allows the verification of Architecture and Territory in relation to

man, natural light, colours, changing shadows and reflections. The hand-drawing is essential to compare the perception of the objects and the transmission by senses. Ultimately, we can think of the sketch as synthesis and transmission of experience and reality to everybody.



Figure 1: D. Larrain, pencil drawing, a.a. 2007/2008

All great designers have drawn freehand subjects from life, and have used this medium of analysis to observe, understand and design. The architecture's notes, the travel notebooks, and the life-sketches constitute a common heritage for past and contemporary architects and engineers. Life drawing is the first and irreplaceable approach instrument to the knowledge of the world we live in. The image performed on the site, in direct contact with the object of study, allows to transcribe information and data deductible from reality, eliminating the superfluous.



Figure 2: D. Silva, pencil and watercolor, a.a. 2007/2008

As the image produced on site might be detailed, there is always a selection process that allows an intellectual highlight of what could matter. The ability to "learn to see" is expressed mainly in grasping the essential elements and characteristics of the object of study, a skill that requires highly selective perception, combined with technical know-how.



Figure 3: P. Sholz, pen and watercolor, a.a. 2007/2008

The hand drawing is the first and irreplaceable instrument of knowledge for a series of works: architectural design and infrastructure design, conservation, construction and urban surveying. The image made on site, in direct contact with the subject to be represented, allows you to write down information and data deducible from reality, eliminating what is not needed for certain purposes. Matter how detailed the image produced on site at the base there is always a selection process that allows you to highlight what matters. The designer is confronted with the reality that surrounds him and with himself, with his powers of observation, analysis, synthesis and with his technical ability. The same subject is represented in many different ways as there are people who represent it. This is because the design is individual analysis of reality, but also because it depends on subjective factors such as theoretical preparation, knowledge of graphic tech-

niques and methods of representation, the use of materials and tools. The basic design for the architect or engineer is one in which the clarity of reading, symbols and graphic conventions take precedence over the quest for natural and illusory effects. The adherence to the true happen following a correct perspective, a graphic rendering without ambiguity, the correct position and proportion of the various elements depicted. Life drawing in architecture is configured as an analytical tool able to answer certain questions and help you to understand an aspect of reality that surrounds us.



Figure 4: L. Bown, pencil drawing, a.a. 2007/2008

The fundamental condition for drawing from life is the direct contact with the architecture or the landscape, the choice of the system of signs capable of translating the character of reality in an image and examination of the forms and proportions to arrive at a graphical summary. The designer must therefore develop their own code of signs, operational steps that allow you to get in tune with reality in order to transcribe the essential characteristics.



Figure 5: A. Pizzola, pencil drawing, a.a. 2007/2008.

In addition, the designer must actively and critically observe the reality show it and simultaneously use the technical knowledge acquired to grasp the shape of things and their arrangement in space as well as assess the tonal differences that give depth and chiaroscuro effects.

The hand drawing is the representation, visualization critical communication; is the result of a complex set of analysis of the forms, the application of geometric concepts, evaluation tonal, character selection, knowledge of graphic techniques and critical consciousness. Few signs, tracks quickly by a great designer, may be more comprehensive than a long description. Very few contour lines, some architectural elements just mentioned, sparse splashes of color, can be a very expressive graphic and full of suggestions. Every designer has to build, over time their own baggage of signs to represent building materials, architectural details, natural features such as rivers, trees, rocks.

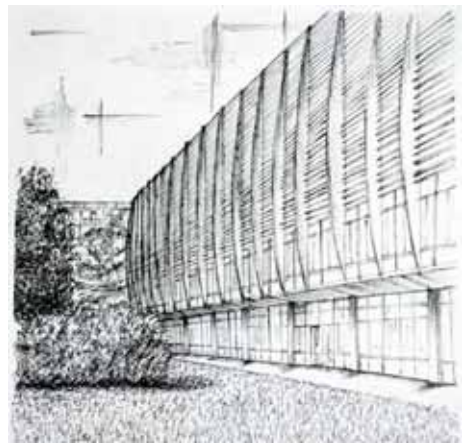


Figure 6: S. Toffano, pencil drawing, a.a. 2006/2007

The life drawing allows us to observe how things change depending on the brightness, the distance from the point of view and proximity to other elements; allows us to understand that the signs are never ends in themselves, that the

technique should not prevail on observation and transcription of meanings. The signs must express concepts and technique should be used to express with simply and clearly the purpose for which it was backed design. The hand has importance to confer expressive quality to the sign and with the hand play an important role the form, the pressure, the speed and the performance of the gesture, ie factors that make a graphic sign beautiful, expressive and personal.



Figure 7: D. Ruiz, pencil and watercolor, a.a. 2007/2008

The representation of reality is filtered by the personality of the performer and moves through the relationship between architecture and the environment, modern architecture and ancient buildings, places of work and study, everyday objects. The use of color, especially watercolor, due to its characteristics of transparency, allows you to keep accurate and legible signs in pen or pencil, giving the design freshness, brightness and countless chromatic

possibilities.

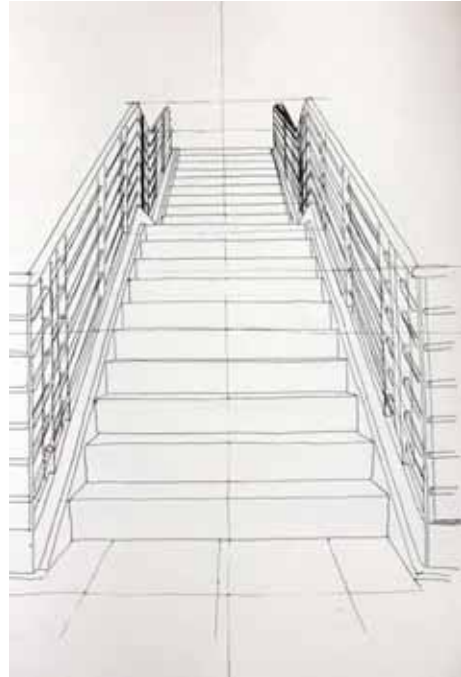


Figure 8: E. Morbelli, pencil drawing, a.a. 2007/2008.

Life drawing is the most complete and profound method to cure, learn and represent Architecture, Landscape, and Environment, ie the world that surround us and the inner self of imagination and creation.

REFERENCES

- [1] A.A.V.V. *Un pensiero per Rosati*. Politecnico di Torino, Torino 1999.
- [2] R. Bizzotto, C. Mezzetti, N. Sardo. *Rappresentazione dalla formazione alla professione*, Edizioni Kappa, Roma, 2009.
- [3] L. Carlevaris (a cura di). *Linee di ricerca nell'area del Disegno. Contributi delle tesi di dottorato in Mostra*. Aracne Editrice, Roma, 2013.



Figure 9: S. Toffano, pen and watercolor, a.a.
2006/2007

ABOUT THE AUTHORS

MARIA PAOLA MARABOTTO, PhD Architect in Architectural Drawing and History and Theory of Representation. She is currently adjunct professor at the School of Architecture at the Polytechnic of Milan in the courses of Representation of the Executive Plan and Architectural Drawing. Among her previous studies include monographs *Gocce di Luce. Fontane in Torino* (2003) and *Tracce, immagini, suggestioni di un territorio. Le vie del sale nel basso Piemonte* (2006). Participate actively in conferences and exhibitions on architectural drawing, surveying and representation.

RULED SURFACES OF CONSTANT SLOPE IN 3-MINKOWSKI SPACE

Željka MILIN ŠIPUŠ¹ and Ljiljana PRIMORAC GAJČIĆ²

¹University of Zagreb, Croatia

²University of Osijek, Croatia

ABSTRACT: In this paper we analyze ruled surfaces of constant slope in a special ambient setting – 3-dimensional Minkowski space. We are particularly interested in ruled surfaces with null rulings, which have no Euclidean counterparts. We also analyze conditions on invariants of a ruled surface to obtain rulings-preserving isometries.

Keywords: Minkowski space, isometry, ruled surface, *B*-scroll

1. INTRODUCTION

A ruled surface is a surface traced out by a straight line moving along a curve c and therefore it admits a local parametrization of the form

$$f(u, v) = c(u) + ve(u), u \in I \subset \mathbf{R}, v \in \mathbf{R}. \quad (1)$$

The curve c is the base curve (the generating curve) and the straight lines with directions $e(u)$ are the rulings of a surface. Alternatively, when normalized, $e(u)$ can be regarded as a spherical curve (the director curve) and $(u, v) \mapsto ve(u)$ as a cone (the director cone). In classical Euclidean differential geometry, ruled surfaces are described, up to Euclidean motion, by the moving Sannia frame (G. Sannia, 1925.) which provides a complete system of the Euclidean invariants of a ruled surface – curvature κ , torsion τ and striction σ .

A special type of ruled surfaces are surfaces of constant slope, that is, surfaces whose rulings make a constant angle with a fixed direction (called the axis of the surface). In Euclidean geometry those surfaces are characterized by the constant conical curvature $k = \frac{\tau}{\kappa} \neq 0$, having, therefore, a cone of revolution for the director cone. Ruled surfaces with conical curvature $k = 0$ are surfaces whose director cone is a plane (conoidal surfaces). In Euclidean geometry, any ruled surface can be bent by rulings-preserving isometries (so called, Minding isometries) to a

surface of constant slope and to a conoidal surface.

Motivated by this, our interest is to analyze analogous properties of ruled surfaces in a different ambient setting – in 3-dimensional Minkowski space \mathbf{R}_1^3 . The space \mathbf{R}_1^3 is a real affine space whose underlying vector space is endowed with a indefinite symmetric bilinear form of index 1 (a pseudoscalar product). Ruled surfaces in 3-dimensional Minkowski space are classified with respect to the casual character of their base curve and their rulings (spacelike, timelike or null (lightlike, isotropic)). We are specially interested in the surfaces with null rulings, so-called class \mathbf{M}_0 . Among them, *B*-scrolls of null Frenet curves are of further interest. We show that rulings of surfaces of constant slope in the class \mathbf{M}_0 share properties with null generalized helices, in particular, with null cubics. Furthermore, we also investigate conditions on invariants of a ruled surface to obtain rulings-preserving isometries of such surfaces and we give some examples concerning surfaces of constant slope.

Finally, it is important to mentioned that besides the class of *ruled* surfaces of constant slope, some other closely related classes of surfaces have been investigated in last years in various ambient settings – surfaces of constant slope and surfaces of constant angle (e.g. [3, 7, 8]). Surfaces of constant slope are surfaces whose

position vector makes a constant angle with the normal at each point whereas surfaces of constant angle are surfaces whose normal vectors make a constant angle with a fixed vector. The first class of these surfaces, surfaces of constant slope, are "the surface analogue" of logarithmic spirals whose characterizing property is that the angle between their tangent lines and position vectors are constant.

2. PRELIMINARIES

Let \mathbf{R}_1^3 be a Lorentz-Minkowski space (a pseudo-Euclidean space), that is, the standard vector and topological space \mathbf{R}^3 equipped with the indefinite symmetric bilinear form (a pseudoscalar product)

$$x \cdot y = x_1 y_1 + x_2 y_2 - x_3 y_3.$$

A vector x in the Lorentz-Minkowski 3-space is called spacelike if $x \cdot x > 0$ or $x = 0$, timelike if $x \cdot x < 0$ and null (lightlike, isotropic) if $x \cdot x = 0$ and $x \neq 0$. The pseudonorm of a vector x is defined as the real number

$$||x|| = \sqrt{|x \cdot x|} \geq 0.$$

The Lorentzian cross-product of vectors x, y is defined by

$$x \times_L y = J(x \times y),$$

where on the right-hand side \times denotes the Euclidean cross product and

$$J = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

In \mathbf{R}_1^3 the following quadrics are introduced

$$\begin{aligned} S_1^2(p, r) &= \{q \in \mathbf{R}_1^3 : (q - p) \cdot (q - p) = r^2\}, \\ H^2(p, r) &= \{q \in \mathbf{R}_1^3 : (q - p) \cdot (q - p) = -r^2\}, \\ LC(p) &= \{q \in \mathbf{R}_1^3 : (q - p) \cdot (q - p) = 0\}. \end{aligned}$$

As surfaces in Euclidean space, these sets appear as a rotational hyperboloid of one-sheet, a

rotational hyperboloid of two-sheets and a rotational cone. In \mathbf{R}_1^3 , the set $S_1^2(p, r)$ is called the Lorentzian sphere or a pseudosphere with center p and radius $r > 0$, the set $H^2(p, r)$ a hyperbolic plane with center p and radius $r > 0$, and $LC(p)$ the light cone with the vertex p . We use notation $S_1^2 = S_1^2(0, 1)$, $H^2 = H^2(0, 1)$, $LC(p) = LC$. The set $S_1^2(p, r)$ inherits a pseudo-Riemannian metric of index 1 from the ambient \mathbf{R}_1^3 (a pseudo-Euclidean metric), $H^2(p, r)$ a Riemannian metric (it appears as the model for the hyperbolic plane, therefore the name), and $LC(p)$ inherits a degenerated metric of rank 1 (an isotropic metric).

3. NULL RULED SURFACES OF CONSTANT SLOPE IN 3-MINKOWSKI SPACE

In Euclidean geometry, a ruled surface S is called a surface of constant slope if its rulings make a constant angle with a fixed direction. Those surfaces are characterized by the constant conical curvature $k = \frac{\tau}{\kappa} \neq 0$, where κ, τ are curvature and torsion of a ruled surface. Ruled surfaces with conical curvature $k = 0$ are surfaces whose director cone is a plane (conoidal surfaces).

We treat the same problem in Minkowski geometry. Ruled surfaces are surfaces which admit a parametrization of the form (1) and are classified with respect to the casual character of their base curve and their rulings (spacelike, timelike or null (lightlike, isotropic)). It can be shown that, after an appropriate change of the base curve, classes can locally be represented either by having the base curve and rulings non-null and linearly independent (the class \mathbf{M}_1^1) or null and linearly independent (the class \mathbf{M}_0) (see [4]). It is of special interest to analyze the latter ones.

Example 3.1. *The class \mathbf{M}_0 includes surfaces such as parabolic null cylinders, which are flat, and pseudo-spheres $S_1^2(p, r)$, which are non-flat and totally umbilical.*

In Minkowski geometry, the angle between vectors when one vector is null (lightlike,

isotropic) is not well-defined. Nevertheless, analogously to the definition of null-curves of constant slope (generalized null helices), we say that a ruled surface with null rulings e is a surface of constant slope if there exists a constant vector $q \neq 0$ such that

$$q \cdot e = \text{const.} \quad (2)$$

To obtain analogous results that characterize null-ruled surfaces of constant slope, first we introduce the analogue of the Sannia frame for a null-ruled surface S , as a Darboux frame of the director cone.

Let S be a non-cylindrical surface, $e' \neq 0$. Because e' is orthogonal to e , e' must be a spacelike vector. Then the (spacelike) director curve $e(u)$ can be parametrized by the arc-length u , $e'(u) \cdot e'(u) = 1$. Furthermore, for the spacelike curve $e(u)$ lying on the lightcone LC , a null basis $\{e, e', y\}$ of \mathbf{R}_1^3 can be introduced (see [6]) so that

$$e \cdot e = y \cdot y = 0, \quad e' \cdot e' = 1, \quad e \cdot y = 1, \quad e \cdot e' = e' \cdot y = 0$$

and the following formulas hold

$$e'' = ke - y, \quad y' = -ke'. \quad (3)$$

The function $k : I \rightarrow \mathbf{R}$ is called the conical curvature of the ruled surface. It is, in fact, the geodesic curvature of a curve $e(u)$ on LC . For arc-length parametrized curves $e(u)$ it is given by

$$k(u) = -\frac{1}{2}e''(u) \cdot e''(u).$$

If we denote by \bar{k} , $\bar{\tau}$ the curvature and the torsion of a spacelike, arc-length parametrized director curve $e(u)$ in \mathbf{R}_1^3 and by $t = e'$ its tangent, n principal normal and b binormal, then the Frenet formulas state

$$t' = \bar{k}n, \quad n' = -\bar{k}t + \bar{\tau}b, \quad b' = -\bar{\tau}n.$$

Furthermore the following relations can be established

$$\bar{k} = \sqrt{-2\epsilon k}, \quad \bar{\tau} = -\frac{1}{2}\left(\frac{k'}{k}\right), \quad \epsilon = n \cdot n. \quad (4)$$

Theorem 3.2. *Let S be a non-cylindrical null-ruled surface and let $e(u)$ be its director curve, $k(u)$ its conical curvature. Then S is a surface of constant slope if and only if*

$$k = \text{const.}$$

Proof. If S is a surface of constant slope, then there is a constant vector $q \neq 0$ such that (2) holds. Let us put $q = q_1e + q_2e' + q_3y$, where $q_1 = q \cdot y$, $q_2 = q \cdot e'$, $q_3 = q \cdot e$. The assumption (2) implies $q_3 = \text{const.}$ Now by differentiating (2) we get $q_2 = q \cdot e' = 0$. Similarly $q'_1 = q \cdot y' = -kq \cdot e' = 0$. Therefore we have $q = q_1e + q_3y$, where $q_1, q_3 = \text{const.}$ Since q is constant, $q' = (q_1 - q_3k)e' = 0$, and S is non-cylindrical, $e' \neq 0$, we have $q_1 - q_3k = 0$, or

$$k = \frac{q_1}{q_3} = \text{const.}$$

Notice that $q_3 \neq 0$, since otherwise $q = q_1e$. This implies $0 = q' = q_1e'$ and therefore $q_1 = 0$, that is, $q = 0$, a contradiction. The vector q which generates the axis of a surface of constant slope can be written as

$$q = q_3(ke + y), \quad q_3 = \text{const.} \neq 0.$$

Conversely, if $k = \text{const.}$, then consider the vector $q = q_3(ke + y)$, $q_3 = \text{const.}$ The vector q is a constant vector, $q' = q_3(ke' + y') = 0$, and it holds $q \cdot e = q_3$. Therefore, S is a surface of constant slope. \square

By sharpening the result from [6], we get:

Corollary 3.3. *Let S be a non-cylindrical null-ruled surface of constant slope and let $e(u)$ be its director curve (with the arc-length parameter u). Then $e(u)$ is a planar curve. More precisely, e is a (pseudo)-circle, i.e. an ellipse in a spacelike plane ($k < 0$), a parabola with the axis parallel to the lightlike direction in a lightlike plane ($k = 0$) or a hyperbola in a timelike plane ($k > 0$).*

Proof. The director curve $e(u)$ is a spacelike curve with Minkowski curvatures \bar{k} , $\bar{\tau}$ related

to k by (4). Theorem 3.2 implies that $\bar{\tau} = 0$, and therefore $e(u)$ is a planar curve. If $k \neq 0$, $e(u)$ lies in its osculating plane spanned by e' , $e'' = ke - y$, i.e. a plane with the Minkowski normal $q = ke + y$. This normal is a spacelike (resp. timelike) vector, and therefore, a plane containing e is timelike (resp. spacelike) if and only if $k > 0$ (resp. $k < 0$). Furthermore, since, $\bar{\kappa} = \text{const.}$, the curve e is a (pseudo)-circle. Notice that the plane Minkowski normal $q = ke + y$ is the axis of S .

When $k = 0$ then $\bar{\kappa} = 0$ and $\bar{\tau}$ is not defined. In this case, $e''(u)^2 = 0$ and $e''(u) = -y(u) \neq 0$, hence e'' is a null (lightlike) vector. Now $y' = 0$ implies $e''' = 0$ and therefore $e(u)$ is a parabola, $e(u) = -\frac{1}{2}y(u)u^2 + a_1u + a_2$, where a_1, a_2 are constant vectors. It lies in a plane spanned by e', y which is lightlike (containing only one lightlike direction – the direction of a null vector y). The axis of a parabola $e(u)$ is the straight line determined by y , therefore $e(u)$ is a parabolic circle. \square

Contrary to the case of Euclidean space, null-ruled surfaces with $k = 0$ are not conoidal. We can notice that $k = 0$ if and only if $q^2 = 0$, that is, if and only if the axis of a ruled surface of constant slope is null (lightlike). Moreover, we have the following:

Corollary 3.4. *Let S be a non-cylindrical null-ruled surface. Then S is never a conoidal surface.*

Proof. Let S be a non-cylindrical null-ruled surface. Let us assume that S is conoidal and let π be its director plane with the Minkowski normal q , $q = q_1e + q_2e' + q_3y$. Notice that since $e(u)$ are parallel to the plane π , π must be either timelike or lightlike. Now $q_3 = e \cdot q = 0$ and $e' \cdot q = e'' \cdot q = 0$. The last equation implies $q_1 = y \cdot q = 0$. Hence q is collinear to e' , $q = q_2(u)e'(u)$. Since $q_2 = e' \cdot q = 0$, we have $q = 0$, a contradiction. \square

3.1 Null-ruled surfaces of constant slope as B -scrolls

Every null scroll satisfying $c' \cdot e \neq 0$ admits a parametrization as a B -scroll, i.e. as a ruled surface with null rulings and with a base curve c which is a null Frenet curve with a null frame (A, B, C) ,

$$c' = A, A^2 = B^2 = 0, C^2 = 1,$$

$$A \cdot B = 1, A \cdot C = 0, B \cdot C = 0.$$

Rulings are determined by the field $e = B$. Therefore, a B -scroll is parametrized by

$$f(u, v) = c(u) + vB(u), \quad u \in I \subset \mathbf{R}, v \in \mathbf{R}.$$

Vector fields (A, B, C) satisfy analogues of the Frenet formulas

$$A' = k_2C, B' = k_3C, C' = -k_3A - k_2B, \quad (5)$$

where functions k_2, k_3 are called the curvatures of c . B -scrolls are timelike surfaces with the first fundamental form I and the shape-operator S given by

$$I = \begin{pmatrix} k_3(u)^2 v^2 & 1 \\ 1 & 0 \end{pmatrix}, \quad (6)$$

$$S = \begin{pmatrix} k_3(u) & 0 \\ k_2(u) + vk_3'(u)^2 & k_3(u) \end{pmatrix}.$$

The Gaussian curvature K and the mean curvature H of a B -scroll are given by

$$K = k_3(u)^2, \quad H = k_3(u).$$

Examples of B -scrolls include flat (and therefore minimal) B -scrolls, such as parabolic null cylinders, and non-flat totally umbilical ones, such as pseudo-spheres $S_1^2(p, r)$.

Example 3.5. *A pseudosphere $S_1^2(p, r)$ is a ruled surface whose both families of rulings are null. The following parametrization is the local parametrization of $S_1^2(0, r)$ as a B -scroll*

$$f(u, v) = (r, 0, 0) + \frac{1}{\sqrt{2}}(0, u, u) +$$

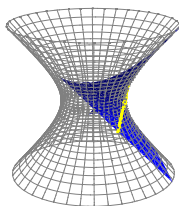


Figure 1: A pseudosphere $S_1^2(p, r)$ with a patch parametrized as a B -scroll and its base curve

$$v\left(-\frac{u}{r}, -\frac{u^2}{2\sqrt{2}r^2} + \frac{1}{\sqrt{2}}, -\frac{u^2}{2\sqrt{2}r^2} - \frac{1}{\sqrt{2}}\right).$$

The director curve $e(u)$ of $S_1^2(p, r)$ is a parabola in the plane $y - z = \sqrt{2}$ with the axis $(0, 1, 1)$ pointing to the lightlike direction of that plane (to the vector A). Its curvatures are $k_2 = 0$, $k_3(u) = -\frac{1}{r}$. Hence it is a surface of constant Gaussian and mean curvatures.

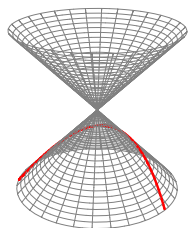


Figure 2: Director curve $e(u)$ of $S_1^2(p, r)$ on the lightlike cone LC

Theorem 3.6. A surface S is a B -scroll of constant slope if and only if:

1. $k_2 \neq 0$ and $\frac{k_3}{k_2} = \text{const.}$. If $\frac{k_3}{k_2} = \text{const} \neq 0$, then S is a B -scroll over a generalized null helix, and if $k_3 = 0$, then S is a B -scroll over a generalized null cubic (then S is a cylindrical surface).

2. $k_2 = 0$. Then S is a B -scroll over a straight null-line, in particular, $S_1^2(p, r)$.

Proof. Let q be a constant vector such that $q \cdot B = \text{const.}$ We can write $q = q_A A + q_B B + q_C C$, where $q_A = q \cdot B = \text{const.}$, $q_B = q \cdot A$, $q_C = q \cdot C$. By differentiating q_A , we get

$$q'_A = q \cdot B' = k_3 q_C = 0$$

and therefore either $k_3 = 0$ or $q_C = 0$.

- If $k_3 = 0$, then c is a generalized null cubic ([4]). We have $B' = 0$, so B is a constant field and S cylindrical. Therefore, S is a B -scroll over a generalized null cubic. Its axis is not unique, it holds $q \cdot B = \text{const.}$ for any constant vector q .
- If $q_C = 0$, then $q'_B = (q \cdot A)' = k_2 q_C = 0$, i.e. q_B is a constant function as well. Moreover, $0 = q'_C = -k_3 q_B - k_2 q_A$.

If $k_2 \neq 0$ then $\frac{k_3}{k_2} = -\frac{q_A}{q_B} = \text{const.}$ This characterizes the base curve c of S as a generalized null helix ([4]). Then S is a B -scroll over a generalized null helix with the axis

$$q = q_A A + q_B B, \quad q_A, q_B = \text{const}$$

Conversely, if $\frac{k_3}{k_2} = \text{const.}$, then the constant vector $q = q_A \left(A - \left(\frac{k_2}{k_3} \right) B \right)$, $q_A = \text{const.} \neq 0$, is the axis of a surface of constant slope, $q \cdot B = q_A$.

If $k_2 = 0$, then $0 = q'_C = -k_3 q_B$ with $k_3 \neq 0$ implies $q_B = 0$ and therefore the axis of S is determined by $q = q_A A$ which is a null vector. Furthermore, since $A' = 0$, the tangent vector $c' = A$ is a constant vector, hence the base curve c is a straight line of null direction. The vector field C satisfies $C' = k_3(u)A$, therefore it can be determined by

$$C(u) = A \left(\int k_3(s) ds \right) + A_1,$$

where A_1 is a constant vector. Since C is a unit spacelike vector, we have $A \cdot A_1 = 0$, $A_1 \cdot A_1 = 1$. Finally,

$$B(u) = \int k_3(s)C(s)ds + B_1,$$

$$B_1 = \text{const.}, B_1 \cdot B_1 = 0.$$

By this, the null-frame of the base curve c is determined. Notice that the director curve $B(u)$ is a (part of a) parabola with an axis A . This can be seen by introducing a new parameter $s^* = \int k_3(s)ds$. Then the curve $C(s^*) = As^* + A_1$ is a straight-line. Since $\frac{ds^*}{ds} = k_3(s)$ and $\frac{dB(s)}{ds} = k_3(s)C(s)$, we have that $B(s^*) = \int C(s^*)ds^*$, and therefore $B(s^*)$ parametrizes a parabola.

Conversely, if $k_2 = 0$, then the constant lightlike vector $q = q_A A$, $q_A = \text{const.}$, generates the axis of a surface of constant slope, $q \cdot B = q_A$.

□

Example 3.7. A curve parametrized by

$$c(s) = (f(s), \frac{1}{\sqrt{2}}(s + \frac{g(s)}{2}), (\frac{1}{\sqrt{2}}(s - \frac{g(s)}{2}))),$$

where f, g are C^2 -functions, $g' = (f')^2$, is a generalized null cubic. Its Frenet frame is given by

$$A(s) = (f'(s), \frac{1}{\sqrt{2}}(1 - \frac{g'(s)}{2}), \frac{1}{\sqrt{2}}(1 + \frac{g'(s)}{2}))$$

$$B(s) = (0, \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}), \quad C(s) = (1, -\frac{f'(s)}{\sqrt{2}}, \frac{f'(s)}{\sqrt{2}})$$

and its curvatures by $k_2(s) = f''(s)$, $k_3(s) = 0$. A B -scroll with over a generalized null cubic is a cylindrical surface, therefore flat and minimal.

Example 3.8. A null-ruled surface with $k_3 = \text{const.} \neq 0$ is called the isoparametric surface ([2]). These surfaces are surfaces of constant Gaussian and mean curvatures. If such a surface is a surface of constant slope, then, when parametrized as a B -scroll, $k_2 = \text{const.}$

as well. A special case is $S_1^2(p, r)$. The general parametrization of such a surface is given in ([2]) by

$$f(u, v) = (g(u), 0, -au + \int g'(u) \cos u du) +$$

$$v(\cos u, \sin u, 1).$$

Example 3.9. If $k_2 = 0$, then S is a surface of constant slope. Figure 3 shows a surface with $k_3(u) = 2u$, hence a surface of constant slope but with non-constant curvatures K and H .

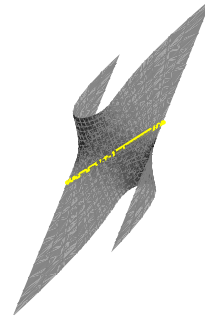


Figure 3: A surface of constant slope with $k_2 = 0$, $k_3(u) = 2u$

3.2 B -scrolls with a base curve parametrized by the pseudoarc-length parameter

Finally, we use special parameters on a base curve c of a B -scroll. A base curve c with $c''(s) \cdot c''(s) > 0$ (hence $k_2 \neq 0$) can be reparametrized by, so called, pseudoarc-length parameter, $c''(s) \cdot c''(s) = 1$. Then

$$c' = A, \quad A' = C, \quad B' = k_L C, \quad C' = -k_L A - B, \quad (7)$$

where the function $k_L = \frac{1}{2}c'''^2$ is called the lightlike curvature of c . It holds that c is a generalized null helix if and only if it is a null helix ([4]), i.e. a curve with constant lightlike curvature. Therefore, the following holds:

Corollary 3.10. *Let S be a B -scroll of constant slope over a null Frenet curve c parametrized by pseudoarc-length parameter. Then c a generalized null helix ($k_L = \text{const.}$) and it is congruent to one of the following, with $k_L = \frac{\sigma^2}{2} > 0$, $k_L = -\frac{\sigma^2}{2} < 0$ or $k_L = 0$, respectively:*

$$\begin{aligned} c_1(s) &= \left(\frac{1}{\sigma^2} \cos(\sigma s), \frac{1}{\sigma^2} \sin(\sigma s), -\frac{s}{\sigma} \right), \\ c_2(s) &= \left(-\frac{s}{\sigma}, \frac{1}{\sigma^2} \cosh(\sigma s), \frac{1}{\sigma^2} \sinh(\sigma s) \right), \\ c_3(s) &= \left(\frac{s^3}{4} - \frac{s}{3}, \frac{s^2}{2}, \frac{s^3}{4} + \frac{s}{3} \right). \end{aligned}$$

By a simple calculation it follows:

Corollary 3.11. *The rulings of a B -scroll of constant slope over a null Frenet curve c parametrized by pseudoarc-length parameter are generated by*

$$\begin{aligned} B_1(s) &= \left(-\frac{\sigma}{2} \sin(\sigma s), \frac{\sigma}{2} \cos(\sigma s), \frac{\sigma}{2} \right), \\ B_2(s) &= \left(-\frac{\sigma}{2}, -\frac{\sigma}{2} \sinh(\sigma s), -\frac{\sigma}{2} \cosh(\sigma s) \right), \\ B_3(s) &= \left(-\frac{3}{2}, 0, -\frac{3}{2} \right). \end{aligned}$$

Proof. We use equations that follow from (7),

$$A = c', \quad C = c'', \quad B = -kA - C = -kc' - c'''. \quad \square$$

Remark 1. Notice that the director curve of a B -scroll generated by B_1 (resp. B_2) is a circle (resp. pseudocircle), in accordance with Corollary 3.3 on the null-ruled non-cylindrical surfaces of constant slope. The B -scroll generated by the base curve c_3 and rulings B_3 is a cylindrical surface, and therefore, not classified in Corollary 3.3. Furthermore, notice that B -scrolls having parabolas as director curves do not appear in the Corollary 3.10, since they have $k_2 = 0$ and their base curves cannot be parametrized by the pseudoarc-length.

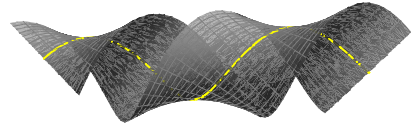


Figure 4: B -scroll of constant slope with the base curve c_1

3.3 Isometries of B -scroll that preserve rulings

Two surfaces are said to be locally isometric if there exists a smooth mapping between them whose differential preserves the pseudoscalar product of their tangent vectors ([1]). Intuitively, these surfaces can be transformed one to another by smooth bending, without tearing or stretching. In Minkowski space, the casual type of vectors is preserved by local isometries, i.e. spacelike (resp. timelike, null) vectors are mapped again to spacelike (resp. timelike, null) vectors. Local isometry also preserves a type of a surface, thus mapping spacelike (resp. timelike, null) into surfaces of the same type, by preserving their first fundamental forms.

Let us consider a local isometry between two ruled surfaces that additionally preserves the rulings of surfaces. Such an isometry is called a Minding isometry of ruled surfaces ([5, 9]). This can be considered only when rulings of surfaces are of the same casual type. In particular, we consider local isometries of B -scrolls (they are time-like surfaces). We can, without loss of generality, assume that the base curve is mapped into a base curve. Therefore, in local coordinates we can take $\bar{u} = u$, $\bar{v} = v$. From the first fundamental form (6) we know that k_3 is preserved (hence K and H as well) and we can vary k_2 without affecting the first fundamental forms. Therefore, by using Theorem 3.6 we can obtain the following analogue of the Euclidean result:

Theorem 3.12. *All B -scrolls are Minding isometric to B -scrolls of constant slope.*

4. CONCLUSIONS

In this paper we have analyzed ruled surfaces of constant slope in a special ambient – 3-dimensional Minkowski space. We were particularly interested in ruled surfaces with null rulings, which have no Euclidean counterparts. We have analyzed them in three different settings: generally, parametrized as null-ruled surfaces, then, parametrized as B -scrolls and finally as B -scrolls with a base curve parametrized by the pseudoarc-length. Our interest in ruled surfaces of constant slope comes from the assertion proved at the end of the paper, that all ruled surfaces with null rulings in 3-dimensional Minkowski space can be bent to null-ruled surfaces of constant slope by rulings-preserving isometries.

REFERENCES

- [1] H. Brauner. *Differentialgeometrie*, Friedr. Vieweg & Sohn, 1981.
- [2] F. Dillen, I. Van de Woestyne, L. Verstraelen, J. Walrave. Ruled surfaces of finite type in 3-dimensional Minkowski space, *Results Math.*, **27**: 250–255, (1995).
- [3] Y. Fu, X. Wang. Classification of Timelike Constant Slope Surfaces in 3-Dimensional Minkowski Space, *Results Math.*, **63**: 3–4, 1095–1108, (2013).
- [4] J. Inoguchi, S. Lee. Null curves in Minkowski 3-space, *Int. Electron. J. Geom.*, **1**: 40–83, (2008).
- [5] E. Kruppa. Natürliche Geometrie der Mindingschen Verbiegungen der Strahlflächen, *Monatsh. Math.*, **55**: 340–345, (1957).
- [6] H. Liu. Curves in the Lightlike Cone, *Beitr. Algebra Geom.* [Contributions to Algebra and Geometry], **45**: 291–303, (2004).
- [7] R. López, M. I. Munteanu. Constant angle surfaces in Minkowski space, *Bull. Belg. Math. Soc. Simon Stevin*, **18**: 271–286, (2011).
- [8] M. I. Munteanu, A.-I. Nistor. A New Approach on Constant Angle Surfaces in \mathbb{E}^3 , *Turk. J. Math.*, **33**: 169–178, (2009)
- [9] H. Pottmann, J. Wallner. *Computational Line Geometry*. Springer, 2001.

ABOUT THE AUTHORS

1. Željka Milin Šipuš is a full professor at the Department of Mathematics, Faculty of Science, University of Zagreb, Croatia. She studied mathematics at University of Zagreb. Her research interests include geometry, in particular differential geometry of curves and surfaces in different ambient spaces, and mathematics education, where she currently studies understanding of mathematical concepts in physics. She can be reached by e-mail: zeljka.milinsipus@math.hr
2. Ljiljana Primorac Gajčić is a research assistant at the Department of Mathematics, University of Osijek. She can be reached by e-mail: lprimorac@mathos.hr

SEJFRIEDIAN: EXISTENCE, UNIQUENESS, CONSTRUCTING AND THE PROOF OF PROPERTIES

Vladimir SHELOMOVSKII

Russia

ABSTRACT: The paper presents the properties of an amazingly beautiful geometrical construction made from five triangles and four circles, called the Sejfriedian in honor of its discoverer Michael Sejfried [1–4]. The definition of the Sejfriedian is given, the existence and uniqueness theorems are proved, two ways of constructing are given and many properties most of which had previously been found numerically by Michael Sejfried are proved. We have created the set of dynamic illustrations using freely distributed DGS GInMA which makes it convenient to introduce the Sejfriedian study in training courses on geometry. This study enables to demonstrate deep connections between various transformations (for example, an inversion and projective transformations [7–9]), the possibilities of flat problems solving with the use of a view from outside (from the space). The Sejfriedian properties research allows to introduce students to little-known concepts of a simedian, the Lemoine point, the Fermat points, the Apollonius points and more [5],[8]. The richness of Sejfriedian properties and possibilities of their use in training courses may be comparable to the richness of the conic family. The paper contains a familiarization version of the research. To see all proves and formulas download the paper full version [6]. Install DGS GInMA [10], and then each figure becomes the key to enter the corresponding interactive drawing.

Keywords: Projective transformation, invariant, inversion, affine property, DGS GInMA.

1. INTRODUCTION

Ten years ago Polish researcher Michael Sejfried opened an exceptionally beautiful construction which has a great potential in terms of geometry [1–3]. Michael Sejfried was spell-bound by its beauty and devoted many years to numerical investigation of its properties with the help of computer programs Geometer's Sketchpad and Mathematica. We name this construction a Sejfriedian in honor of Michael Sejfried. In the early stages of research the construction consisted of three triangles and a circle. In this paper the definition of the Sejfriedian is given, the existence and uniqueness theorems are proved, two ways of constructing are given and many properties are proved which had been previously found numerically by Michael Sejfried [1], [2]. We have created the set of dynamic geometric illustrations using

freely distributed DGS GInMA [10] which makes it convenient to introduce the Sejfriedian study in training courses on geometry.

2. DEFINITIONS AND DENOMINATIONS

We use the Sejfried notation for the Sejfriedian definition. Let an arbitrary triangle KLM and the point K_1 be given. The triangle vertices are named in clockwise direction. The point K_1 lies on the arc KL of the KLM circumcircle ω centered at the point O .

We call $K_1L_1M_1$ the *Sejfried amicable triangle* for KLM if it is inscribed into ω and there is a projective transformation mapping ω into a circle, $K_1L_1M_1$ and KLM into regular triangles. We call the triangles KLM and $K_1L_1M_1$ the *Sejfried pair*, ω is called the *Sejfried circle*.

Let the point A be the intersection point of the straight lines KM and K_1L_1 , B be the inter-

section point of KL and M_1L_1 , C be the intersection point of LM and K_1M_1 . The triangle ABC have been called the *Sejfried generating triangle*.

We name the tangent AA_L “left” if the acute angle clockwise rotation map AA_L onto AO . Let the left tangents of ABC vertices touch with ω at the points A_L , B_L , and C_L , and the right tangents touch with ω at A_R , B_R , and C_R . We call the triangles $A_LB_LC_L$ and $A_RB_RC_R$ the *Sejfried generated triangles*.

Let the circles ω_A , ω_B , and ω_C are perpendicular to ω and are centered at ABC vertices with radii $R_A = AA_L$, $R_B = BB_L$, $R_C = CC_L$. We name ω_A , ω_B , ω_C the *Sejfried satellite circles*.

The construction formed by five Sejfried triangles (the Sejfried pair, the generating triangle and two generated triangles) and four circles (the Sejfried circle and three satellite circles) is named *the Sejfriedian*.

We call the *Sejfried number* the ratio

$$Se = \frac{KK_1 \cdot M_1L_1 \cdot ML}{K_1M_1 \cdot L_1M \cdot LK}.$$

Let the point S be the intersection point of the straight lines KM and BC , T be the intersection point of KL and AC , U be the intersection point of LM and AB . The parameter *stu*, used by Sejfried, is determined by the formula

$$stu = \frac{AU \cdot BS \cdot CT}{BU \cdot CS \cdot AT}.$$

The *Sejfried function* is a function (in the general case it is two-dimensional function) allowing to identify the position of the point O with respect to ABC .

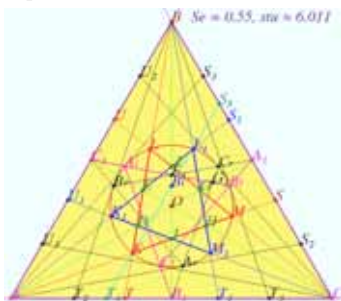


Figure 1: Regular Sejfriedian.

If the Sejfried pair consists of the regular triangles then the Sejfriedian have been called the *regular Sejfriedian*. The example of the regular Sejfriedian is shown in Figure 1.

If the Sejfried circle ω is inscribed into the generating triangle ABC , the Sejfriedian have been called the *base Sejfriedian*. The example of the base Sejfriedian and Sejfriedian in general form are shown in Figures 2 and 3.

The Sejfried pair is shown in red and blue, the generating triangle is highlighted by yellow, the generated triangles are shown in black and green, the Sejfried circle is shown in crimson, the satellite circles are shown in black.



Figure 2: Base Sejfriedian.

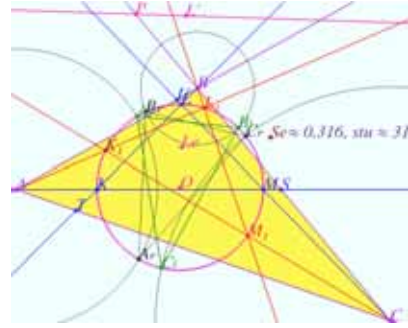


Figure 3: Sejfriedian in general form.

3. MAIN THEOREMS

3.1 Existence theorem

There exists at least one Sejfriedian for an arbitrary triangle KLM and the point K_1 on the arc KL of the KLM circumcircle.

There exists at least one Sejfriedian for an arbitrary triangle KLM and the given Sejfried number $Se \in (0,1]$.

3.2 Uniqueness theorem

There exists one and only one Sejfriedian for an arbitrary triangle KLM and the point K_1 on the arc KL of the KLM circumcircle if the Lemoine (Grabe) point Le of KLM does not coincide with the KLM circumcenter O .

There exists one and only one Sejfriedian for an arbitrary triangle KLM and the number $Se \in (0,1]$ if the Lemoine point Le of KLM does not coincide with the KLM circumcenter O .

3.3 Proved theorems and Corollaries

1. Let an arbitrary triangle and the circumcircle be given, and the circumcircle center O coincides with the Lemoine (Grabe) point Le of the triangle. Then the triangle is regular.

1.a. Each pair of regular triangles inscribed in one circle is the Sejfried pair.

2. Let an arbitrary triangle and the incircle be given, and the incenter I coincides with the Gergonne point G . Then the triangle is regular.

3. Any two regular triangles inscribed in one circle form the Sejfried pair.

3.a. Let the projective transformation map the incircle of the given triangle into a circle. Then it map the preimage Gergonne point into the image Gergonne point, the Gergonne point is an invariant of this transformation.

4. Let the projective transformation map the incircle of the given triangle into a circle and map the Gergonne point into the circle image center. Then the given triangle is mapped into the circumscribed regular triangle.

5. Let the projective transformation map the circumcircle of the given triangle into a circle. Then the preimage Lemoine point is mapped into the image Lemoine point, the Lemoine point is an invariant of such transformation.

5.a. Let the projective transformation map the circumcircle of the given triangle into a circle and the Lemoine point into the image circumcenter. Then the given triangle is mapped into the regular triangle, inscribed into the circle image.

6. Let the projective transformation map the Sejfried circle into a circle. Then the Sejfried pair is mapped into a Sejfried pair.

7. Let the projective transformation map the Sejfried circle into a circle and map the Lemoine point into the image circumcenter. Then the Sejfried pair is mapped into the Sejfried pair of regular triangles.

7.a. Any Sejfriedian is affinely equivalent to the regular Sejfriedian under the projective transformation which map the Sejfried circle into a circle.

7.b. The polar line of the common Lemoine point for the Sejfried pair with respect to the Sejfried circle is the preimage of the line of infinity of the regular Sejfriedian.

7.c. The preimages of parallel straight lines (which may be constructed for any regular Sejfried pair) intersect at the polar line of the common Lemoine point with respect to the Sejfried circle.

7.d. The preimages of any group of points equidistant from the common Lemoine point of the regular Sejfriedian lie on a conic.

8. The straight line BC , the straight lines KK_1 , LM_1 , ML_1 and the straight lines $A_R A_L$, $B_R C_L$, $B_L C_R$ are concurrent. The point of concurrence belongs to the polar line of the Lemoine point with respect to the Sejfried circle.

8.a. The straight lines AC , KM_1 , MK_1 , LL_1 , $A_R C_L$, $B_R B_L$, $A_L C_R$ intersect at one point (point E in Figure 4). The straight lines AB , LK_1 , KL_1 , MM_1 , $A_R B_L$, $B_R A_L$, $C_L C_R$ intersect at one point (point D in Figure 4). The points of intersection P , D and E belong to the polar line of the Lemoine point with respect to the Sejfried circle.

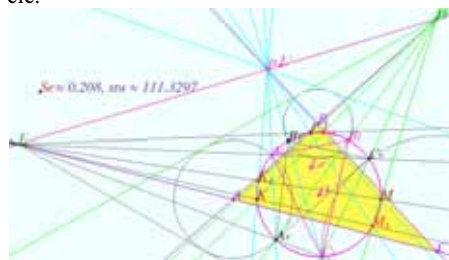


Figure 4: Three groups of collinear lines.

4. AUXILIARY CONSTRUCTIONS

4.1 Perspector construction for the given circle

Let us construct the perspector S of the projective transformation, which map the circle ω centered at O into the circle ω' , and the given point L , different from O , into the center of ω' .

Construction Let the radius OL be intersected by the circle ω at the point L_s , the point L_0 be diametrically opposite to L_s , the segment SL_s be perpendicular to the circle ω plane, and the ratio holds (see Figure 5):

$$\frac{SL_s}{OL_s} = \sqrt{\frac{2LL_s}{OL}}.$$

In this case the perpendicular L_sS_1 from the vertex of the right angle L_s to SL_0 is halved by the segment connecting L and S . The point O_1 is the midpoint of L_sS_1 .

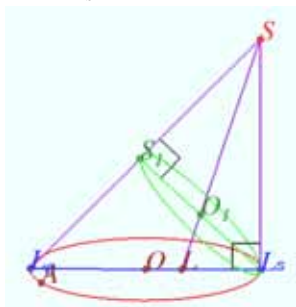


Figure 5: Perspector S constructing.

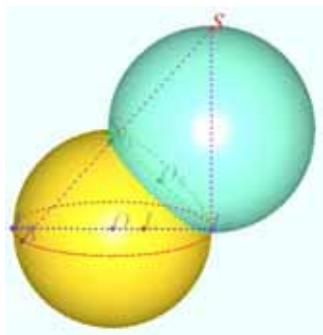


Figure 6: Stereographic projection mapping a circle into a circle.

We construct the sphere with the diameter SL_s . It is shown in Figure 6 by blue. The sphere with the diameter L_sL_0 is shown by yellow. The circle ABC is the big circle of the constructed yellow sphere. The intersection curve of the spheres is the circle ω' . According to the definition, the circle ω is the stereographic projection of the circle ω' under the projection from the perspector S onto the plane ALO .

4.2 Construction of the projective transformation for an isosceles triangle

Let us consider the projective transformation mapping an isosceles triangle into the regular triangle. Figure 7 shows some of the possible situations.

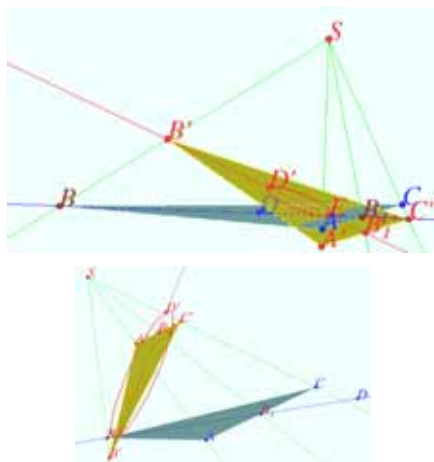


Figure 7: Isosceles triangle mapped into the regular triangle.

The equilateral triangle is shown by green. The isosceles triangle is shown by blue. It has the vertex B with the angle equal to 2β . The perspector of the transformation is the point S . The intersection line of the triangles planes passes through one of the ends of the Seifried circle diameter DE , which lies on the isosceles triangle axis of symmetry. If the angle B is less than 60° , the point D is closer to the vertex B , than E . If the angle B is greater than 60° , the point E is closer to the vertex B , than D .

Let us consider the view in the direction

along AC . Figure 8 shows two possibilities of the planes position. The red line corresponds to the equilateral triangle $A'B'C'$ axis of symmetry, B'_1 is the midpoint of its base. The blue line corresponds to the isosceles triangle ABC axis of symmetry, B_1 is the midpoint of its base. The angle BEE' between the triangles planes is denoted α . We use the parameter τ :

$$\tau = \frac{2h'}{3R'} = \frac{2b'}{R'\sqrt{3}} = \sqrt{\frac{s+1}{s-2}}, s = Se + \frac{1}{Se},$$

where h' is the height, R' is the Seifried circle radius, b' is the half AC side.

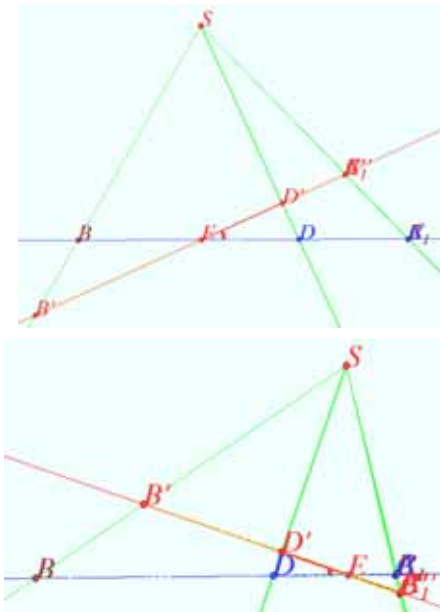


Figure 8: View along AC .

1. Let the apex B angle be less than 60° . This case is shown in the upper drawing of Figure 8. The ratio of the line $B'B'_1$ segments is:

$$B'D' : D'E : EB'_1 = \frac{2h'}{3} - R' : 2R' : \frac{h'}{3} + R'.$$

We use the view in the direction along BB_1 and find the ratio, using the similar triangles SAB_1 and $SA'B'_1$:

$$\frac{BB_1}{B'B'_1} = \frac{\tan BSE + \tan ESB_1}{\cos \alpha \cdot (\tan B'SD' + \tan B'_1SD')}.$$

$$\frac{BB_1}{B'B'_1} = \frac{1}{\sqrt{3} \cdot \tan \beta \cdot \cos^2 \alpha \cdot (1 + \tau_1 \cdot \tan^2 \alpha)}.$$

We obtain the equation for the angle α :

$$\cos \alpha = \frac{\sqrt{3} \cdot \tan \beta}{\tau + 1} + \sqrt{\frac{3 \cdot \tan^2 \beta}{(\tau + 1)^2} + \frac{\tau - 1}{\tau + 1}}.$$

2. Let the apex B angle be more than 60° . This case is shown in the bottom drawing of Figure 8:

$$\cos \alpha = -\frac{\sqrt{3} \cdot \tan \beta}{\tau - 1} + \sqrt{\frac{3 \cdot \tan^2 \beta}{(\tau - 1)^2} + \frac{\tau + 1}{\tau - 1}}.$$

We construct the Seifriedian knowing the angle α and the ratio of $BD : DE : EB_1$.

4.3 Construction of the triangle with the given circumcenter and the Lemoine point

Let us construct the triangle ABC such that the vertex A , the circumcenter O and the Lemoine point L be given. If the points A , O , and L are concurrent, we decide that the plane ABC is known.

Construction We construct step by step the circle $\omega(O, OA)$, the Brocar axis LO , the points of intersection of the circle ω and LO (the points L_1 and L_0), the perspector S and the circle image ω' .

We find the point A_1 as a result of intersection of the straight line AS and the circle ω' .

We construct the regular triangle $A_1B_1C_1$ inscribed into the circle ω' . It is shown in the left drawing of Figure 9. The blue triangle $A_1B_1C_1$ is inscribed into the green circle ω' .

We fulfill the projective transformation of the triangle $A_1B_1C_1$ with the perspector S onto the plane AOL . This is shown in the right drawing of Figure 9. The purple rays from the perspector S cross the red circle ω at the points A , B and C . The triangle ABC is constructed.

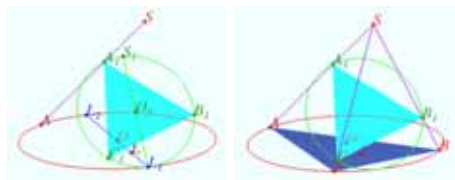


Figure 9: Image $A_1B_1C_1$ and preimage ABC .

4.4 Amicable Sejfried triangle construction

Let an arbitrary triangle KLM and the point K_1 belonging to the arc KL of the circle ω (KLM) be given. Use a projective transformation to construct a Sejfried triangle.

Construction We construct the perspector S and the circle ω' . Figure 10 shows the construction of the triangle KLM image noted as $K'L'M'$ and the point K_1 image noted as K_1' on the circle ω' .

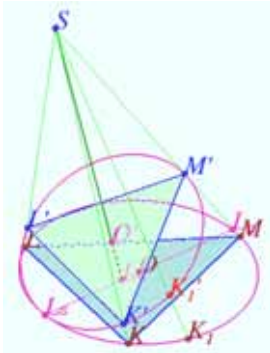


Figure 10: Constructing of the images $K'L'M'$ and K_1' .

We construct the regular triangle $K_1L_1M_1$ inscribed in a circle ω' (Figure 11).

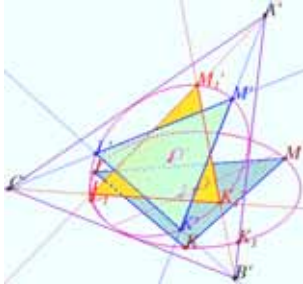


Figure 11. Constructing of the images $K_1'L_1'M_1'$ and $A'B'C'$

We find the vertices of the regular generating triangle $A'B'C'$ at the intersections of the straight lines which contain the sides of the triangle. We construct the Sejfriedian preimages $K_1L_1M_1$ and ABC using the corresponding points of the Sejfriedian images $K_1'L_1'M_1'$ and $A'B'C'$ (see Figure 12).



Figure 12: Constructing of $K_1L_1M_1$.

4.5 Sejfriedian construction with the use of one vertex of the amicable triangle

Let an arbitrary triangle KLM and the point K_1 belonging to the arc KL of the circle ω (KLM) be given. Construct a Sejfriedian.

Construction We construct the Brocar axis OLe and the polar line of the Lemoine point Le . The polar line is perpendicular to the Brocar axis and contains the point L' inverted to Le with respect to the circle ω . $OLe \cdot OL' = OK^2$ (see Figure 13).

We find the point P as a result of intersection of KK_1 and the polar line.

We construct the vertices L_1 and M_1 of the Sejfried triangle as the second points of intersection of the straight lines PL and PM with the circle ω (L and M are the first points).



Figure 13: Sejfriedian construction with the use of one vertex of the amicable triangle

We construct the generating triangle ABC using the intersection points of the straight lines containing the sides of the Sejfried triangles.

We construct the generated Sejfried pair using the tangents.

4.6 Sejfriedian construction with the use of the Sejfried number

Let an arbitrary triangle KLM and the real number $Se, 1 \leq Se > 0$ be given. Construct a Sejfriedian.

Construction We construct the point K_1 according to the condition $\frac{KK_1}{LK_1} = Se \frac{KM}{LM}$. We use the Apollonius circle and the intersection point of this circle with the circumcircle KLM (see Figure 14). Then we use the construction 4.5 of this paper.

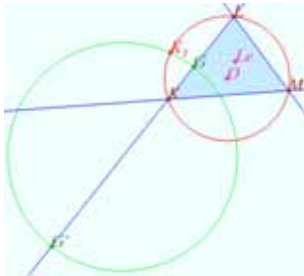


Figure 14: Point K_1 construction.

5. PROPERTIES OF THE SPECIAL TYPES SEJFRIEDIANS

5.1 Regular Sejfriedian properties

The base regular Sejfriedian is shown in Figure 15, the regular Sejfriedian is shown in Figure 16. The point Q is the center of the triangle ABC . The straight lines described by Sejfried, the family of concentric circles uniting the points equidistant from the center Q , the family of straight lines parallel to the base AC , the Sejfried number Se and the parameter stu are shown.

For the base regular Sejfriedian we have $AC_1 = KL = K_1 L_1 = \frac{AB}{2}$, and AC_1 touches the incircle. By the property of a tangent and a secant, several segments meet the conditions of the golden section:

$$\frac{AK_1}{AC_1} = \frac{AC_1}{AL_1} = \frac{AC_1}{AK_1 + AC_1} = \frac{K_1 C_1}{C_1 L_1}.$$

For an arbitrary regular Sejfriedian we take

the generating triangle side length as a unit and denote

$$AT = AU_1 = \dots = \xi \leq \frac{1}{2}.$$

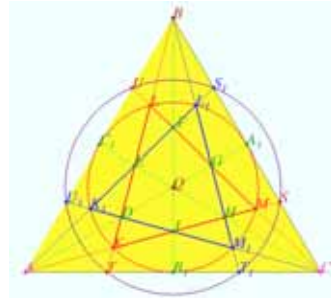


Figure 15: Base regular Sejfriedian.

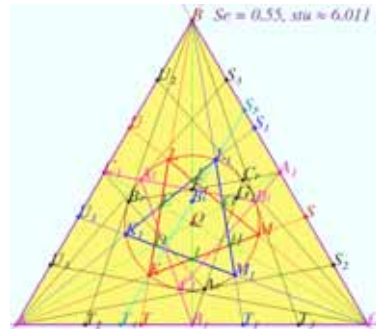


Figure 16: Regular Sejfriedian.

We have

$$AB_1 = AC_1 = \frac{1}{2}, AQ = \frac{1}{\sqrt{3}}, QA_1 = \frac{1}{2\sqrt{3}},$$

$$AS = \eta = \sqrt{1 - \xi + \xi^2} = \frac{\sqrt{1 + Se + Se^2}}{1 + Se}.$$

We use the Sejfried number

$$Se = \frac{\xi}{1 - \xi}, \xi = \frac{Se}{1 + Se}.$$

We use the ratios for skew segments and obtain the length of each element of the Sejfriedian. For example:

$$KL = LM = \xi = \frac{1 - Se}{\sqrt{1 + Se + Se^2}}.$$

$$R_{Se} = QK = \frac{1 - Se}{\sqrt{3(1 + Se + Se^2)}}.$$

According to M. Sejfried, we denote the

points of contact of the straight lines passed from the generating triangle vertices with the Sejfried circle as $A_L, A_R, B_L, B_R, C_L,$ and C_R . We get:

$$AA_L = AA_R = \dots = \sqrt{AK \cdot AM} = \frac{\sqrt{\xi(1-\xi)}}{\eta},$$

Let us denote the intersection point of the straight lines L_1K and AC as T_4 . Then:

$$\frac{CT_4}{AT_4} = (CT_4, B_1A) = \frac{CT_4}{AT_4} \cdot \frac{AB_1}{CB_1} = Se + \frac{1}{Se}.$$

The straight line L_1K is parallel to the side AB and intersects the axis of symmetry BB_1 at the point B_i . B_i is the inverse point for B with respect to the Sejfried circle.

Straight lines containing the vertices of generated triangles. We have the following statement for an arbitrary regular Sejfriedian. For each group of the triples of points $A_1B_1C_L, A_1C_RB_R, B_1C_LA_L, B_1A_RC_R, B_1A_LB_L,$ and $C_1B_RA_R$ the points in the group are concurrent.

Proof Sejfried generated triangles are regular triangles. Since all their defining points are mapped into each other under the rotation turning the triangle ABC by 120° around the point Q , the triangle center. Let us consider the triple of the points $A_1B_1C_L$. The radius QB_L divides the angle $A_1B_1C_L$ in half, the angle $\angle QB_LC_L = 30^\circ$.

The points A_1 and B_L belong to the circle with the diameter BQ . So the sum of the angles is $\angle QB_LA_1 + \angle QBA_1 = 180^\circ$, where $\angle QBA_1 = 30^\circ$. Hence $\angle A_1B_1C_L = \angle QB_LC_L + \angle QB_LA_1 = 180^\circ$ is a flat angle, and the points A_1 and B_L belong to the straight line A_1C_L . We have a similar situation for the other triples of points.

The collinear points. Let us consider an arbitrary regular Sejfriedian. Let the straight lines A_1C_R and C_1B_L intersect at the point F_1 , the straight lines B_1C_R and C_1B_L intersect at the point G_1 . Then F_1 belongs to the median BQ , and G_1 belongs to the median AQ .

Corollary. Let the point A_1 is the intersection point of the straight lines AQ and BC , the point B_1 is the intersection point of the straight lines BQ and AC , the point C_1 is the intersection point of the straight lines CQ and AB . Then the triangle $A_1B_1C_1$ is the generating triangle for

the Sejfried generated triangles.

For an arbitrary regular Sejfriedian we have:

$$\frac{B_L C_R}{B_R C_L} = \frac{|\sqrt{Se} - 1 + Se|}{\sqrt{Se} + 1 - Se}.$$

Ellipses of the Sejfriedian. There are the ellipses which are tangent to the sides of the generated triangle ABC at the points $(U, S_1,$ and $B_1), (S, T_1,$ and $C_1), (T, U_1,$ and $A_1)$. In the case of $Se = \frac{1}{2}$ these ellipses are tangent to the Sejfried circle at the vertices of the Sejfried triangles. One of the ellipses is shown in Figure 17.

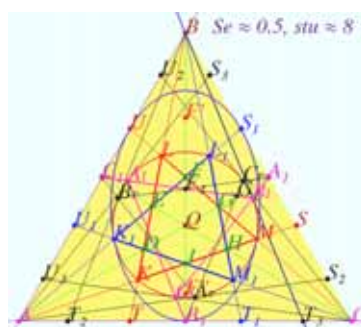


Figure 17: Ellipse in special case $Se = 0.5$.

5.2 Sejfriedian for the isosceles triangle.

Radii of the accompanying circles.

Let the isosceles triangle ABC ($AB = BC$) be given. We want to find the radii of the accompanying circles $R_A = R_C$ and R_B .

Let B_i be the inverse point for B with respect to ω . Let G be the mutual inversion (and dilation) center of the circles centered at A and B . The points G, B_i, K, L_1 and T_4 are collinear.

We prove that the point G belongs to KL_1 . Let us consider the group of the straight lines $AB, M_1L, KL_1,$ and MK_1 . Under the projective transformation mapping the Sejfriedian of the isosceles triangle into the regular Sejfriedian the images of these straight lines are parallel. So they have a common point at infinity, which corresponds to the dilation center of the circles centered at A and B . These circles are perpendicular to ω .

We know that the point T_4 belongs to KL_1 . A

projective transformation retains the double ratio $(CT_4, B_1A) = \frac{CT_4}{AT_4} \cdot \frac{AB_1}{CB_1}$.

For the isosceles triangle the point B_1 is the midpoint of BC . It means that the ratio of any two segments which make up the double ratio with segments AB_1 and CB_1 is saved. In particular

$$\frac{CT_4}{AT_4} = Se + \frac{1}{Se}.$$

We prove that the point Bi belongs to KL_1 . The circles ω and ω_B are perpendicular, the point Bi is the intersection point of ω and ω_B radical axis with ω and ω_B center line. So under the inversion with respect to ω the point Bi is mapped into B , and under the inversion with respect to ω_B it is mapped into O . Under the inversion with respect to ω the straight line KL_1 is mapped into the circle IKL_1 , containing the point B . This follows from the symmetry of the points K and K_1 with respect to the axis of symmetry of the triangle ABC and the symmetry of the points K_1 and L_1 with respect to ω_B . The straight line KL_1 is the radical axis of the circles IKL_1 and ω_B . Hence the angles L_1BI and L_1KI are equal one to another.

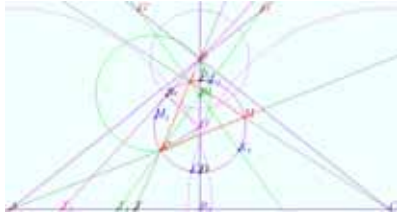


Figure 18. The isosceles triangle Sejfriedian and the straight line KL_1 .

The Menelaus condition for the straight line $GKBi$ which intersects the sides of the triangle ABB_1 leads to the equation:

$$R_B \cdot R_C = BT_2 \cdot R_C + R_C^2 - \frac{AC^2}{2}.$$

The condition that the center O lies on the radical axis of the circles ω_B and ω_C is written in the form of:

$$2R_B \cdot BT_2 = R_B^2 - R_C^2 + BC^2.$$

The solution of the system has the form of:

$$R_C = \frac{AC^2}{2\sqrt{BT_2^2 + AC^2 - BC^2}}.$$

$$R_B = BT_2 - \frac{b^2 + d^2}{\sqrt{d^2 + 3b^2}} = \sqrt{h^2 + d^2} - \frac{b(s-1)}{\sqrt{s+1}}.$$

Isosceles triangle Sejfriedian construction. Let the point B_1 be the midpoint of AC , $b = AB_1$, $h = BB_1$. Using the points T and T_2 belonging to AC we have:

$$\frac{AT}{CT} = \frac{r_A Se}{r_C} = Se,$$

$$\frac{AT_2}{CT_2} = \frac{r_A \tilde{Se}}{r_C}, \tilde{Se} = \frac{\sqrt{Se} - 1 + Se}{\sqrt{Se} + 1 - Se}.$$

$$\frac{B_1 T_2}{AB_1} = \frac{d}{b} = \frac{1 - Se}{\sqrt{Se}}, BT_2^2 = h^2 + d^2, d = B_1 T_2.$$

We express the radii using b , d and Se :

$$R_A = \frac{2b^2}{\sqrt{d^2 + 3b^2}} = \frac{2b}{\sqrt{\frac{1}{Se} + Se + 1}} = \frac{2b}{\sqrt{s+1}}.$$

$$R_B = BT_2 - \frac{b^2 + d^2}{\sqrt{d^2 + 3b^2}} = \sqrt{h^2 + d^2} - \frac{b(s-1)}{\sqrt{s+1}}.$$

We consistently construct the next lines and points (see Figure 18): the circles ω_A and ω_B , the point O as the intersection point of the radical axis of the circles ω_A and ω_B with the straight line BB_1 , the straight line BT_2 , the perpendicular OB_r on BT_2 , the radius OB_r of the circle ω , the straight line BT and the points K and L of the intersection of BT with ω .

Sejfried function. In the case of the isosceles triangle the point O belongs to the axis of symmetry of the triangle ABC . Hence it is a scalar function. We may use the distance BO between the center O and the vertex B . It is evidence that

$$\frac{BO}{h} = 1 + \frac{d^2}{h^2} - \frac{b}{h} \sqrt{1 + \frac{d^2}{h^2} \frac{s-1}{s+1}}.$$

Let us verify the formulae in trivial cases.

1. Let the radius of the Sejfried circle be near zero. In this case $Se \rightarrow 1$, $s \rightarrow 2$,

$$\lim_{s \rightarrow 2} \frac{BO}{h} = F(1) = 1 - \frac{b}{h\sqrt{3}}.$$

It is the Fermat point.

2. Let the Sejfried circle be an incircle. In

this case $Se = \frac{3-\sqrt{5}}{2} \approx 0.3819, s=3$.

$$\lim_{s \rightarrow 3} \frac{BO}{h} = 1 + \frac{b^2}{h^2} - \frac{b}{h} \sqrt{1 + \frac{b^2}{h^2}}.$$

It is the incenter.

3. Let the Sejfried circle be near the circumcircle. In this case $Se \rightarrow 0, s \rightarrow \square$,

$$\lim_{s \rightarrow \infty} \frac{BO}{h} = F(0) = \frac{1}{2} + \frac{b^2}{2h^2}.$$

It is the circumcenter.

In Figure 19 the Sejfried function

$\frac{BO}{h}(Se)$ for seven values of $\frac{b}{h} = \left\{ \frac{3}{2}, \frac{5}{4}, 1, \frac{1}{\sqrt{3}}, \frac{1}{2}, \frac{1}{4}, \frac{1}{7} \right\}$ is shown. For

the regular triangle $\frac{b}{h} = \frac{1}{\sqrt{3}}$, the point O is stationary and located at the center of the triangle ABC , $\frac{BO}{h} = \frac{2}{3}$. For an arbitrary triangle the Sejfried function changes monotonically from $F(0)$ to $F(1)$. If $\beta < 30^\circ$, then $\frac{b}{h} < \frac{1}{\sqrt{3}}$.

The function increases monotonically with increasing Se from $F(0)$ to $F(1)$. The point O moves away from the vertex B and approaches to the base AC .

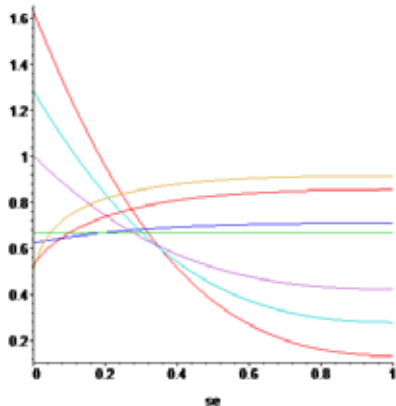


Figure 19: Sejfried function for the isosceles triangle.

Let the angle in the vertex B be more than

60° , $\frac{b}{h} > \frac{1}{\sqrt{3}}$. In this case, the function decreases monotonically with increasing value of the number Se from $F(0)$ to $F(1)$.

5.3 Base Sejfriedian construction

Let the triangle ABC with $AB \neq BC$ be given. Let us construct the base Sejfriedian.

Constructing. Let $\omega = A_1B_1C_1$ be the ABC incircle ($A_1B_1C_1$ be the points of contact). We construct the point D as the intersection point of the lines A_1C_1 and AC , the point E as the intersection point of the lines A_1B_1 and CC_1 (see Figure 20). The points M and K_1 are the points of intersection of DE and ω (see Figure 21). These points are the vertices of two Sejfried triangles. Construction of the remaining vertices is obvious.



Figure 20: Building of the base Sejfriedian



Figure 21: Building of the base Sejfriedian

5.4 Sejfriedian in the case of $Se = 0,5$.

Let the Sejfried number be equal to 0,5. It is trivial to prove that for the regular Sejfriedian the points A_1, L_1 and U are collinear, and that the straight line A_1L_1 is tangent to the Sejfried circle. Similarly the triples of points $(A_1, T_1$ and $M)$, $(B_1, M_1$ and $S)$, $(B_1, U_1$ and $K)$, $(C_1, K_1$ and $T)$, $(C_1, S_1$ and $L)$ have such properties. It is shown in Figure 22 for the regular Sejfriedian and in Figure 23 for an arbitrary one.

The straight lines S_1T, T_1U, U_1S contain the point O (in this point the straight lines A_1A, B_1B, C_1C are intersecting).

The straight lines A_1C_1 , BT_1 , TS_1 are concurrent. Similarly the triples of lines (A_1C_1, BT, T_1U) , (A_1B_1, SU_1, CU) , (A_1B_1, CU_1, ST) , (B_1C_1, T_1U_1, AS) , and (B_1C_1, SU_1, AS_1) are concurrent.

The straight lines A_1C_1 , BB_1 , CU , and AS_1 are concurrent. Similarly the lines A_1B_1 , CC_1 , AS , BT_1 and the lines B_1C_1 , AA_1 , BT , CU_1 are concurrent. These properties are invariant under projective transformations so every Sejfriedian has such properties if $Se = 0,5$.

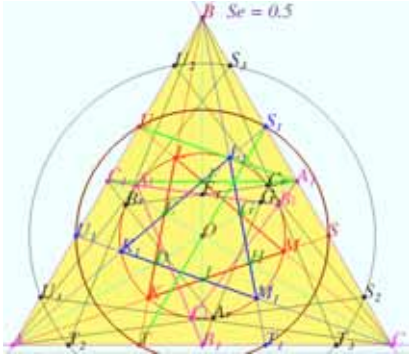


Figure 22: Regular Sejfriedian for $Se = 0,5$.

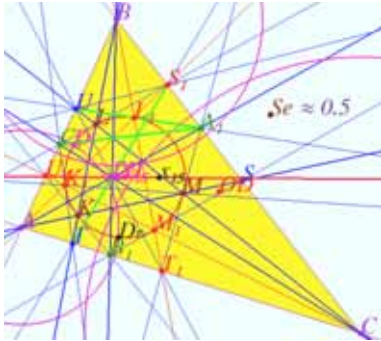


Figure 23: Sejfriedian for $Se = 0,5$.

5.5 Sejfriedian in the case of $Se = 0,618...$

Let the Sejfried number be equal to

$$Se = \frac{\sqrt{5}-1}{2} \approx 0.618. \text{ It is trivial to prove that}$$

for the regular Sejfriedian the points S_1 , U_1 and L , are concurrent and the straight line U_1L is tangent to the Sejfried circle. Similarly the triples of points (U, L_1, S) , (S_1, T_1, M) , $(S, M_1,$

$T)$, (T_1, U_1, K) , and (U, K_1, T) have such properties. It is shown on Figure 24 for the regular Sejfriedian and on Figure 25 for an arbitrary one.

The polar line A_LA_R of the vertex A contains the points T_1 , U , the intersection point of the lines AS_1 , C_1C and BT , and the intersection point of the lines AS , CU_1 and BB_1 . Similarly the polar line B_LB_R contains the points U_1 , S , the intersection point of the lines AS_1 , C_1C and BT , and the intersection point of the lines CU , BT_1 and AA_1 . The polar C_LC_R contain the points S_1 , T , the intersection point of the lines AS_1 , C_1C and BT , and the intersection point of the lines CU , BT_1 and AA_1 .

The points A_1 , C_1 , L_1 and L are collinear. Similarly the points $(A_1, B_1, M_1$ and $M)$, $(B_1, C_1, K_1$ and $K)$ are collinear.

These properties are invariant under projective transformations so every Sejfriedian has such properties.

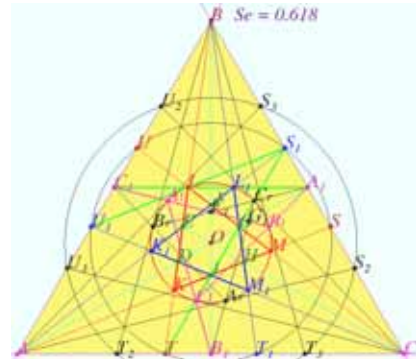


Figure 24: Regular Sejfriedian for $Se = 0,618...$

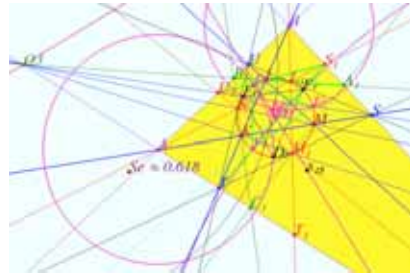


Figure 25: Sejfriedian for $Se = 0,618...$

6. SOME GEOMETRIC FACTS

6.1 Invariants of projective transformations

Four points ratio. Let a circle and belonging to it points A, B, C , and D be given. We call a t_2 -ratio of four points the ratio:

$$t_2 = \frac{|A-B| \cdot |C-D|}{|B-D| \cdot |C-A|}.$$

Let the projective transformation map the circle ABC into a circle and the given points are mapped into the points A', B', C' , and D' , respectively. Then the t_2 -ratio is an invariant under this transformation:

$$t_2 = \frac{|A-B| \cdot |C-D|}{|B-D| \cdot |C-A|} = \frac{|A'-B'| \cdot |C'-D'|}{|B'-D'| \cdot |C'-A'|}.$$

2n points ratio. Let a circle and belonging to it $2n$ points A, B, C, D, \dots, Z be given. We call a t_n -ratio of $2n$ points the ratio

$$t_n = \frac{|A-B| \cdot |C-D| \cdot \dots}{|B-D| \cdot \dots \cdot |Z-A|}.$$

Let the projective transformation map the circle ABC into a circle and the given points are mapped into the points $A', B', C', D', \dots, Z'$, respectively. Then the t_n -ratio is an invariant under this transformation:

$$t'_n = \frac{|A'-B'| \cdot |C'-D'| \cdot \dots}{|B'-D'| \cdot \dots \cdot |Z'-A'|}, t_n = t'_n.$$

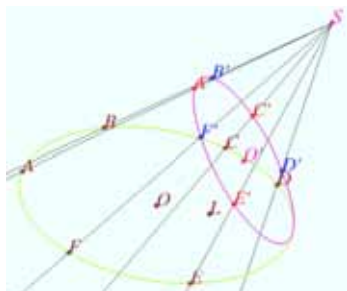


Figure 26: Projective transformation.

Clearly about the ratios. We may formulate the statement 6.1 of this paper in another way. Let us take a circle and inscribe in it a closed broken line, passing once through each given point. Then we paint the segments of the broken line through one in two colors, such as

red and blue as it is shown in Figure 27. We find the product of the lengths of blue segments and the product of the lengths of the red segments. The ratio of this products does not change under the projective transformation mapping the given circle into any other circle. We see the calculated value of $t - t'$ at the top of Figure 27.

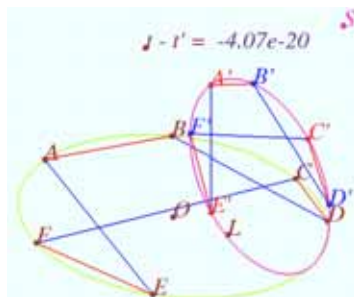


Figure 27: Illustration of invariance.

6.2 Tangents proportion

Let the circle, the points A and B lying on it and the tangents AA_R, AA_L, BB_R , and BB_L be given. We use the subscript « R » for the tangent AA_R if the ray AO is mapped into the ray AA_R under the acute angle clockwise rotation. Let the straight line $A_R B_R$ intersect the straight line

AB at the point E . Then: $\frac{AA_R}{BB_R} = \frac{AE}{BE}$.

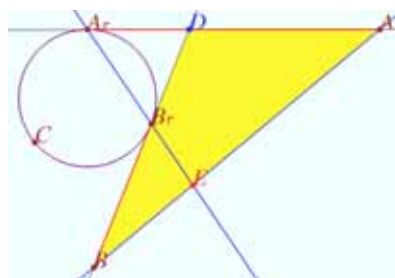


Figure 28: Tangents proportion.

Proof Let the straight line AA_R intersects the line BB_R at the point D , $DB_R = DA_R$. In accordance with the Menelaus' theorem for the triangle ABD , we have:

$$DA_R \cdot AE \cdot BB_R = DB_R \cdot BE \cdot AA_R \Leftrightarrow \frac{AA_R}{BB_R} = \frac{AE}{BE}.$$

6.3 The Newton-Gauss line

The Newton-Gauss line of the cyclic quadrangle contains the incenter.

6.4 The property of the intersecting tangents

Let the circle, the points A and B lying on it and the tangents AA_R , AA_L , BB_R , and BB_L be given. The intersection points of the tangents are C , C' , D , D' , as it is shown in Figures 29 and 30. Then the following statements are valid:

- the straight lines CC' , A_LB_L and A_RB_R are concurrent, the point E of intersection belongs to AB ,
- the straight lines DD' , A_LB_R and A_RB_L are concurrent, the point E' of intersection belongs to AB ,
- the tangents from the point E touch the circle at the points lying on DD' ,
- the tangents from the point E' touch the circle at the points lying on CC' ,
- the straight lines AB_R and BA_L intersect at the point C_1 belonging to CC' ,
- the straight lines AB_L and BA_L intersect at the point D_1 belonging to DD' ,
- the straight lines AB_R and BA_R intersect at the point D_2 belonging to DD' ,
- the straight lines AB_L and BA_R intersect at the point C_0 belonging to CC' .

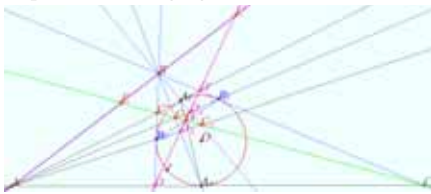


Figure 29: Property of the intersecting tangents.

Under the conditions 6.4, $DD \square OE$, $CC \square OE$, $AB \square OG$, the triangle EGE' is a self-polar triangle, the point O is the orthocenter of the triangle $EE'G$, the point G is the orthocenter of the triangle $EE'O$.

Under the conditions 6.4 the radical axis of the circles (A, AA_R) and (B, BB_R) is perpendicular to AB . The radical axis contains the center

O and the point G . If the circles intersect, the radical axis contains the points H and H' of their intersection.

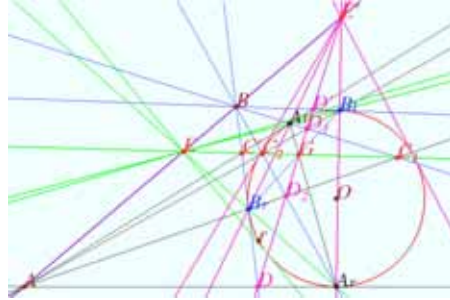


Figure 30: Property of the tangents.

Under the conditions 6.4 the Newton-Gauss line of the cyclic quadrangle $CDC'D'$ contains the center O of the inscribed circle ($AM = BM$, $DM_1 = D'M_1$, $CM_2 = C'M_2$).

Under the conditions 6.4 the following equalities hold:

$$\frac{AE}{AA_R} = \frac{BE}{BB_R}, \frac{AE'}{AA_R} = \frac{BE'}{BB_R}.$$

7. SEJFRIEDIAN PROPERTIES

7.1 Collinearity

Regular Sejfriedian image center. Let Q be the intersection point of straight lines AA_1 and BB_1 . Then Q lies on the straight line CC_1 and is the common Lemoine point of the Sejfried triangles.

Proof. The point Q is the preimage of the regular triangle center under the projective transformation mapping a Sejfriedian into the regular Sejfriedian. The Lemoine point is an invariant of the transformation.

Set of points. The points from the set of points, collinearity of which have been indicated by Sejfried [1-3], for example, A , D , Q , G , G_1 , A_1 , are really collinear.

Proof The collinearity is trivial for the regular Sejfriedian. All Sejfriedians are affine equivalent. Collinearity is the affine property.

Concurrence. The straight lines LL_1 , A_1C_1 , A_LC_R , EG , SU_1 , K_1M , DH , ..., AC are parallel or

intersect at one point belonging to the polar line of the Lemoine point Q with respect to the Sejfried circle.

Proof The statements are trivial for the regular Sejfriedian. All Sejfriedians are affine equivalent. The statements are affine properties. The polar line of the point Q with respect to the Sejfried circle is on the line of infinity.

Straight lines containing the pairs of vertices of Sejfried triangles. The straight lines BC , KK_1 , LM_1 , ML_1 , A_LA_R , B_LC_R , and C_LB_R are parallel or intersect at one point belonging to the polar line of the Lemoine point with respect to the Sejfried circle.

Proof The statements are trivial for the regular Sejfriedian. All Sejfriedians are affine equivalent (3.3).

Tangents. The straight lines SM_1 , TK_1 , UL_1 , LS_1 , MT_1 , KU_1 are tangent to the Sejfried circle.

Tangents passed from the vertices of the generating triangle. The tangents to the Sejfried circle passed from the vertices of the generating triangle are intersecting at the points belonging to the straight lines which contain the point Q . The tangents to the Sejfried circle passed from the vertices A and B are intersecting at the point belonging to CQC_1 , the tangents passed from the vertices B and C are intersecting at the point belonging to AQA_1 , the tangents passed from the vertices A and C are intersecting at the point belonging to BQB_1 .

Point of concurrence. The tangent to the Sejfried circle DD_R is concurrent with the straight lines EF_R and FE_R . Similarly the straight lines DF_R , EE_R and FD_R , DE_R , ED_R and FF_R , DD_L , EF_L and FE_L , DF_L , EE_L and FD_L , DE_L , ED_L and FF_L are concurrent.

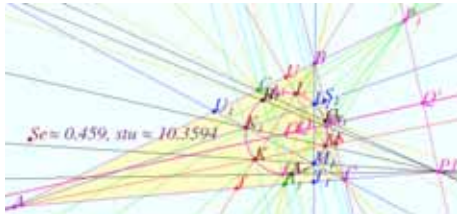


Figure 31: Common point of the straight lines joining the Sejfriedian vertices.

Proof The statements are trivial for the regular Sejfriedian. All Sejfriedians are affine equivalent. The statements are affine properties. The polar line of the point Q with respect to the Sejfried circle is on the line of infinity.

7.2 Invariants

The Sejfried number Se and the parameter stu are invariant under the projective transformation mapping the Sejfriedian into the regular Sejfriedian.

The following ratios hold for any Sejfriedian:

$$\frac{KM \cdot LL_1 \cdot K_1M_1}{KM_1 \cdot L_1K_1 \cdot ML} = \frac{\xi}{1-\xi} = Se = \frac{1}{\sqrt[3]{stu}}, \quad (1)$$

$$\frac{KK_1 \cdot ML_1 \cdot LM_1}{LL_1 \cdot MK_1 \cdot KM_1} = \frac{MM_1 \cdot LK_1 \cdot KL_1}{LL_1 \cdot MK_1 \cdot KM_1} = 1, \quad (2)$$

$$\frac{KK_1 \cdot LL_1 \cdot MM_1}{KM_1 \cdot LK_1 \cdot ML_1} = \left(\frac{\xi}{1-\xi} \right)^3 = Se^3 = \frac{1}{stu}. \quad (3)$$

$$\frac{KA_R \cdot MC_R \cdot LA_R}{MA_R \cdot LC_R \cdot KA_R} = \left(\frac{\xi}{1-\xi} \right)^{1.5} = Se^{1.5}. \quad (4)$$

Proof We use the triple ratios for the sets of following points: K, M, L, L_1, K_1, M_1 for (1), K, K_1, M, L_1, L, M_1 and M, M_1, L, K_1, K, L_1 for (2), K, K_1, L, L_1, M, M_1 for (3), K, A_R, M, C_R, L, A_R and $K_1, A_L, L_1, B_L, M_1, C_L$ for (4). All these points belong to the Sejfried circle. Hence the triple ratios are invariant under the projective transformation mapping the Sejfriedian into the regular Sejfriedian. The equations are trivial for the regular Sejfriedian.

The following ratios hold for any Sejfriedian:

$$\frac{AK \cdot BL \cdot CM}{AA_R \cdot BB_R \cdot CC_R} = \left(\frac{\xi}{1-\xi} \right)^{1.5} = Se^{1.5}. \quad (5)$$

$$AK \cdot BL \cdot CM = AK_1 \cdot BL_1 \cdot CM_1. \quad (6)$$

Proof According to the properties of a tangent and a secant we get six equations:

$$\frac{AK}{AA_R} = \frac{KA_R}{MA_R}, \quad \frac{BL}{BB_R} = \frac{MC_R}{LC_R}, \dots$$

We multiply them by three and find the expression for the segments with endpoints at the Sejfried circle in regular Sejfriedian. Thus we get (5). Since the left and the right tangents

from one point are equal, we obtain (6).

The following ratios hold for any Sejfriedian:

$$\frac{AT \cdot BU \cdot CS}{AU \cdot BS \cdot CT} = \frac{AU_1 \cdot BS_1 \cdot CT_1}{AT_1 \cdot BU_1 \cdot CS_1} = Se^3,$$

$$\frac{AT_2 \cdot BU_2 \cdot CS_2}{AU_2 \cdot BS_2 \cdot CT_2} = \frac{AU_3 \cdot BS_3 \cdot CT_3}{AT_3 \cdot BU_3 \cdot CS_3} = \tilde{Se}^3,$$

Proof The points S, S_1, T, T_1, U, U_1 (and the points $S_2, S_3, T_2, T_3, U_2, U_3$) are concyclic in a regular Sejfriedian, hence they belong to the same quadric in an arbitrary Sejfriedian. According to the Carnot's theorem we get:

$$\frac{AT}{CT} \cdot \frac{AT_1}{CT_1} \cdot \frac{CS}{BS} \cdot \frac{CS_1}{BS_1} \cdot \frac{BU}{AU} \cdot \frac{BU_1}{AU_1} = 1,$$

$$\frac{AT_2}{CT_2} \cdot \frac{AT_3}{CT_3} \cdot \frac{CS_2}{BS_2} \cdot \frac{CS_3}{BS_3} \cdot \frac{BU_2}{AU_2} \cdot \frac{BU_3}{AU_3} = 1.$$

Then we use the double cross-ratios from 7.3.

The following ratio holds for any base Sejfriedian:

$$\frac{AT \cdot BU \cdot CS}{AU \cdot BS \cdot CT} = Se^3 = \left(\frac{\sqrt{5}-1}{2} \right)^6 = \left(\frac{3-\sqrt{5}}{2} \right)^3.$$

7.3 Sejfriedian's double sets

It is known that double cross-ratios are invariant under the projective transformation. So we find the double cross-ratios for the regular Sejfriedian and use them for any Sejfriedian. For example, we write the ratio of a pair of double cross-ratios, which is also an invariant:

$$\frac{(AT, T_1C)}{(AT_1, TC)} = \frac{AT \cdot CT_1}{AT_1 \cdot CT} = \left(\frac{\xi}{1-\xi} \right)^2 = Se^2.$$

$$\frac{CS \cdot BS_1}{CS_1 \cdot BS} = \frac{BU \cdot AU_1}{BU_1 \cdot AU} = Se^2 = stu^{-\frac{2}{3}}.$$

$$(AS, KM) = (BT, LK) = (CU, LM) = Se^3.$$

The following ratios hold for any Sejfriedian:

$$\frac{AK \cdot MS}{r_A \cdot SM_1} = \frac{BL \cdot KT}{r_B \cdot TK_1} = \frac{CM \cdot LU}{r_C \cdot UL_1} = Se^3.$$

$$\frac{A_1B}{r_B} = \frac{A_1C}{r_C}, \frac{B_1A}{r_A} = \frac{B_1C}{r_C}, \frac{C_1A}{r_A} = \frac{C_1B}{r_B}.$$

The following ratios hold for any Sejfriedian:

$$\frac{CS \cdot BA_1}{BS \cdot CA_1} = \frac{BU \cdot AC_1}{AU \cdot BC_1} = \frac{AT \cdot CB_1}{CT \cdot AB_1} = Se,$$

$$\frac{BS_1 \cdot CA_1}{CS_1 \cdot BA_1} = \frac{AU_1 \cdot BC_1}{BU_1 \cdot AC_1} = \frac{CT_1 \cdot AB_1}{AT_1 \cdot CB_1} = Se.$$

$$\frac{CS}{BS} = \frac{CA_1}{BA_1}, Se = Se \frac{r_C}{r_B}, \quad \frac{KK_1}{LK_1} = Se \cdot \frac{KM}{LM}.$$

The point Q of a Sejfriedian has the barycentric coordinates with respect to the triangle ABC equal to:

$$Q = \left(\frac{1}{r_A}, \frac{1}{r_B}, \frac{1}{r_C} \right).$$

The points K, L, \dots of a Sejfriedian have the barycentric coordinates with respect to the triangle ABC equal to:

$$K = \left(\frac{1}{r_A Se}, \frac{Se}{r_B}, \frac{1}{r_C} \right), L = \left(\frac{1}{r_A}, \frac{1}{r_B Se}, \frac{Se}{r_C} \right),$$

$$M = \left(\frac{Se}{r_A}, \frac{1}{r_B}, \frac{1}{r_C Se} \right), K_1 = \left(\frac{1}{r_A Se}, \frac{1}{r_B}, \frac{Se}{r_C} \right),$$

$$L_1 = \left(\frac{Se}{r_A}, \frac{1}{r_B Se}, \frac{1}{r_C} \right), M_1 = \left(\frac{1}{r_A}, \frac{Se}{r_B}, \frac{1}{r_C Se} \right).$$

$$A_1 = \left(0, \frac{1}{r_B}, \frac{1}{r_C} \right), S = \left(0, \frac{1}{r_B}, \frac{1}{r_C Se} \right),$$

Proof The segments AS_1 and BT_1 are inter-

secting at the point L_1 . We use $\frac{BS_1}{CS_1} = \frac{r_B Se}{r_C}$

$$\frac{AT_1}{CT_1} = \frac{r_A}{r_C Se}$$

and the Cheva's formulas for intersecting segments.

7.4 Sejfriedian inverse properties

Let the circles centered at the vertices of the generating triangle and perpendicular to the Sejfried circle be given. Then the vertices of the triangles of the Sejfried pair are mutually inverse with respect to the circles.

Proof The radius of the circle centered at A and perpendicular to the Sejfried circle is equal to $r_A = AA_E$. According to the property of a tangent and a secant we obtain $AK \cdot AM = r_A^2$.

8. CONCLUSIONS

We have proved all Sejfriedian properties described by M. Sejfried in the papers [1–3]. To date, only the Sejfried function, which determines the position of the Sejfried circle center with respect to the generating triangle is not found. The Sejfriedian is a beautiful construction with the variety of easily provable properties. It deserves entering the "golden fund" of the world geometry and worth examining at in-depth study of geometry for the development of students geometric vision. Using geometric freeware DGS GinMA we have created interactive supporting files that illustrate Sejfriedian properties.

REFERENCES

- [1] M.Sejfried. *Amicable triangles and perfect circles.*, at the Proceedings of the 15th International Conference on Geometry and Grafics (ICGG 2012), pp 682-687, 2012. ISBN 978-0-7717-0717-9.
- [2] M.Sejfried. *Amicable triangles and perfect circles.*, Journal for Geometry and Graphics, Volume 17 (2013), No. 1, 53-67.
- [3] M.Sejfried. *Trójkąty zaprzyjaźnione i niektóre ich własności geometryczne.*, Częstochowa, 2010.
- [4] M. Sejfried, V.Shelomovskii. *Elementary Proof of Sejfriedian Properties.*, at the Proceedings of the 17th Asian Technology Conference in Mathematics (ATCM 2012), pp 342-352, 2012. ISBN 978-0-9821164-4-9.
- [5] Yiu P. *Introduction to the Geometry of the Triangle.* Chapter 10. General conics. [<http://www.math.fau.edu/yiu/geometry.html>]
- [6] <http://deoma-cmd.ru/files/documents/Sejfriedian.pdf>
- [7] А. Акопян, А. Заславский. Геометрические свойства кривых второго порядка. М.: МЦНМО, 2007.
- [8] Я. П. Понарин. Элементарная геометрия. Т.1. – М.: МЦНМО, 2004. – 312 с.
- [9] В. Прасолов, В.Тихомиров. Геометрия. Т.1. – М.: МЦНМО, 2013. – 336 с.

Software Packages

- [10] GinMA Software, <http://deoma-cmd.ru/en/Products/Geometry/GInMA.aspx>

ABOUT THE AUTHOR

1. Vladimir Shelomovskii is Associate Professor, Ph.d in mathematics, General Director and leading specialist in developing of math e-books, Deoma LLC. Twenty years he has devoted to high-tech technology, led the research group of the Russian Academy of Sciences engaged in the development of aerodynamic windows for gas lasers. For twenty five years he is working in education as the math lecturer, the Chair of department of mathematical modeling and math methods in economics at the Murmansk University. His research interests include solving of difficult mathematical problems, logic, DGS development and geometry teaching with its use.

A SELF CHECKING CAD TOOL FOR MECHANICAL DRAWINGS IN INTRODUCTORY COURSES

Ichiro TANAKA

Tokyo Denki University, Japan

ABSTRACT: In order to reduce the checking time by teachers in introductory courses for technical drawings, an automated checking system was developed on a CAD system. With this system, students can check their own drawings for themselves, whether their drawings have correct geometry. The system eliminates partitioning differences of graphical elements, recognizes viewpoints and extracts graphical elements in each view of the students' drawings, and compares them with the teacher's model drawing. The checking result is shown by pointing out the improper elements in the students' drawings and warning messages which tell the number of missing elements in each view. The quality of students' drawings improved considerably by adopting this system in introductory courses for mechanical drawings, though total checking time by teachers remained the same because it seems to be still in a saturated situation.

Keywords: Technical drawing education, CAD, Automated drawing checking.

1. INTRODUCTION

Typical exercises in introductory courses for mechanical drawings include the dead copy of sample drawings and the drawings of standard parts of different size from the sample. The main objectives of such exercises include to help students understand the basics of mechanical drawings such as multi-view orthographic projection, line type conventions, cross sectioning, dimensioning rules, etc.

The checking of those drawings by students seems to be easy because the resulting drawings should look exactly the same as the model drawings created by teachers. But in fact, some detailed and precise investigation is required which is difficult to be performed on hand-drawn drawings or printed CAD drawings.

1.1 Typical mistakes in beginners' drawings

To draw a precise sketch, the reference dimensions should not be used to determine the geometry, because they are not to define a necessary dimension but are determined automati-

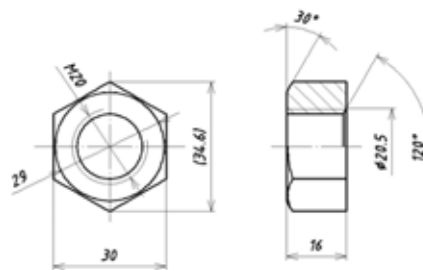


Figure 1 Drawing of a nut

cally by other dimensions. For example, in figure 1, the width across flats (30mm) determines the size of the hexagon, and the reference dimension (the width across corners, 34.6mm) does not indicate the precise value. But in introductory courses, a considerable number of students use this reference dimension (34.6mm) to draw their sketches, which can be regarded as a sign of misunderstanding of drawing rules. Such mistakes are difficult to find out because the error is in sub-mm order.

Figure 2 shows another typical mistake. In this example, the edge between the shaft hole

and the keyway in the cross section is drawn at the position of the imaginary upper border of the shaft hole. The position of this edge should correspond to the corner between the keyway and the hole in the right side view, as indicated by the imaginary line.

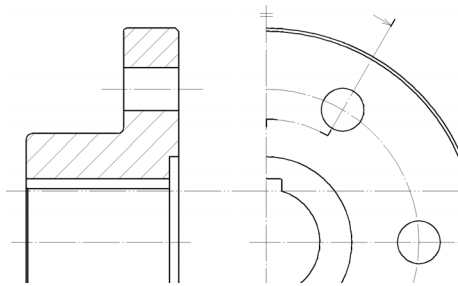


Figure 2 Improper keyway drawing

1.2 Aim of this work

Since 1995, we have introduced CAD for mechanical drawing courses. The use of CAD makes it possible to create precise drawings, and to modify easily when correction is required. Consequently, a long queue of students emerged, waiting for checking by teachers. This is because, in contrast to hand drawing, it is easier to ask teachers to point out their mistakes and modify their drawings than to present perfect drawings at the first trial. Many students try several times.

In order to reduce the waiting time and to improve the quality of students' drawings, we introduced a precise investigation program which runs on the CAD system and which enables students check their own drawings for themselves.

1.3 Background

In 1995, we switched from hand drawing to CAD for the mechanical drawing courses. The main objective of this change was to reduce the instruction time and the load for students. According to the preceding research, students could draw about three times faster with CAD than with drafters, because it took more time to

acquire hand drawing skills.

We customized the CAD (AutoCAD Mechanical) so that students could create the drawing as simply as possible. We prepared layers which correspond to line types such as visible lines, hidden lines and center lines. The color of a line represents the width of the line: cyan for wide lines and white for narrow lines. The dimensions and annotations are customized so that they follow the color convention above, but they are stored in predefined layer (AM_5).

In introductory courses in which the checking system is used, only simple parts and assemblies are drawn. So we assume that only line segments, circles and arcs appear in the drawings.

In our courses, students first use 3D-CAD for basic graphical training, before they use 2D-CAD for designing. So the users of the checking system are familiar with the drawings created by 3D-CAD.

2. PREVIOUS WORKS

There are some commercial tools that detect the differences between two drawings. But they mainly aim to compare different versions of the same drawing, and do not consider about the layout difference.

We found two previous attempt to introduce automated checking system for technical drawing education [1-2]. Both of them developed technical drawing checking systems independent from CAD system, which compare teachers' model drawings and students' drawings. Use of such independent system is not straight forward, because students have to export drawing data in compatible format and then import to the checking system. We are also worried that showing the model drawings might spoil students' interest, for example, to consult standard documents and to determine dimensions properly.

3. CHECKING SYSTEM

3.1 System overview

The checking system is implemented on Au-

toCAD Mechanical with AutoLISP. Figure 3 shows the outline of the checking program. The system consists of two modules, the “student module” and the “teacher module”. The teacher module is used by teachers to create the model data files from the model drawings. These files are delivered to students so that they can check the difference between their own drawings and the models, using the student module.

Both the student module and the teacher module extract graphical elements in the drawings by regularizing the drawings, extracting viewports, and extracting graphical elements in each viewport. This process is described in the following sections.

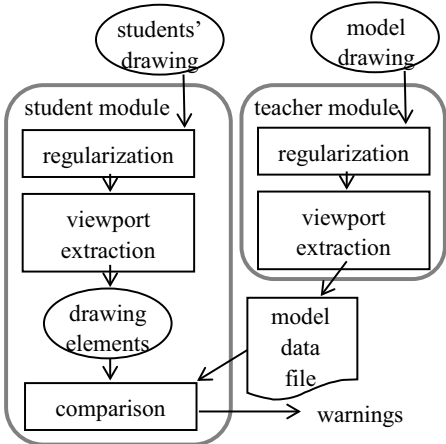


Figure 3 System Structure

3.2 Regularization

Figure 4(a) is an example of technical drawing created by a 3D-CAD. In this drawing, the vertical line on the right side of the nut consists of four line segments (1 to 4). Each of them corresponds to the edge with the same number in the figure 4(b).

On the other hand, it is most likely that this vertical line is drawn as a single line segment in a hand-drawn drawing. Some students may draw the upper-half (cross section) and the lower half (front view) separately. In case of

2D-CAD drawing, there is no concrete rule how this line should be divided (or not).

In order to eliminate such difference, we apply “regularization” process. In this process, connected elements on an identical curve with same line type are merged into one. Actually, co-linear line segments and arcs on the same circle with overlaps are merged into a single line segment or arc.

This regularization guarantees that the element compositions of drawings are same if the outlooks are same.

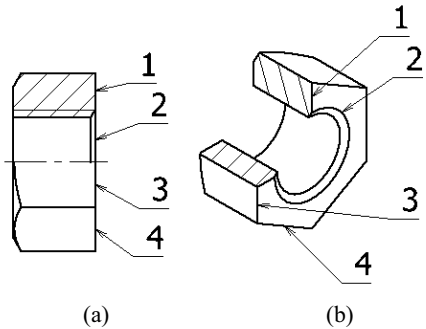


Figure 4 Drawing created by 3D-CAD

3.3 Viewport extraction

The layout of views (viewports) in student’s drawing may vary, and it would not be the same as the teacher’s model drawing. To compare the student’s drawing with the model drawing, the position of each view in the student’s drawing must be identified.

Nagasaka et. al.[2] achieved this by sorting graphical elements. This strategy should be valid if the drawing contains views of single part or assembly, because the relative position between views should meet certain rules. But this may not apply for drawings in introductory courses. Some (but small number of) students ignore the rule about the relative position between views. There are also some exercises to draw more than one part (e.g. a bolt and a nut) in one sheet of drawing.

To extract viewports without any assumption, we use the bounding-box (enveloping rec-

tangle) of each graphical element. Figure 5 shows the viewport extraction process. For each visible line element in the drawing (a), a bounding-box is generated as indicated by dashed-lines in (b). Please note that the bounding-boxes of horizontal lines have zero height and are not seen. Then, connecting or overlapping bounding-boxes are merged until no connecting or overlapping pair exists. The remaining bounding-boxes correspond to the viewports (c). This ensures that the elements inside of other elements (such as the inner circle in the top view) are merged, which will be grouped independently by simply grouping connected elements.

Extracting viewports without any assumption has a merit that checking can be performed before all the viewports are completed.

Once the viewports are extracted, the graphical elements in each viewport as well as the elements that go through the viewport are extracted, so that elements such as center lines and section lines are included. Dimensions and other annotations (stored in layer AM_5) are excluded.

The coordinates of graphical elements are transformed onto the local coordinate system whose origin is the center of the viewport bounding-box, so that they can be compared directly with the elements of model drawings, regardless of the viewport position.

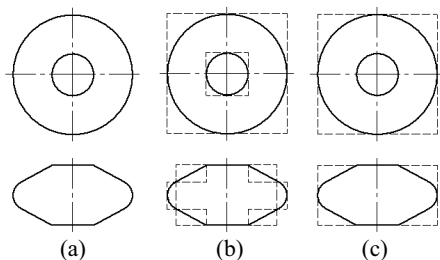


Figure 5 Bounding-box and viewport

3.4 Model data file

A model data file consists of a header and a body. Following information is described in the

header part.

- number of viewports
- viewport names
- whether to regularize or not
- allowable error ratio
- comparison level

The last two elements are mentioned in the next section.

In the body part, the viewports and their graphical elements are described. Each viewport data consists of the following information.

- number of graphical elements in the view
- viewport bounding-box
- list of graphical element data

Each graphical element is described as the list of following.

- element (curve) type (Line, Arc, Circle)
- layer-name (Visible, Hidden, Center, etc., which represents the line type and width)
- geometry data (coordinate of end/center points, radius, etc.)

In the current implementation, the content of a model data file is a text data. We consider encoding them because they have enough information to reproduce the model drawings.

3.5 Comparison process and warnings

After extracting the graphical elements in each view of students' drawings, the system (student module) compares them with the elements in the model data.

For the elements in each view of the student's drawing, the elements in each view in teacher's model drawing is compared, and the number of matching elements is counted. The score for each combination is calculated by the number of matching elements minus the number of remaining and missing elements. The combination with the highest score is adopted as corresponding views.

Once the corresponding view in the teacher's model drawing is decided, warnings are displayed in the following manner.

(1) Warnings for improper elements

The elements in the student's view which are not found in the model view are displayed

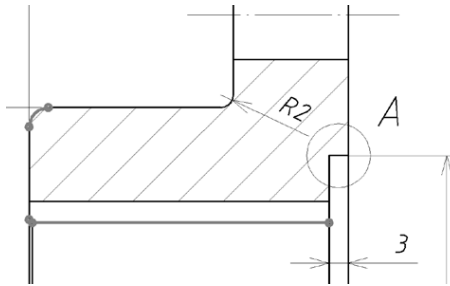


Figure 6 Warnings for improper elements

in magenta, emphasizing the endpoints with circles. In figure 6, improper elements are shown in gray wide lines with circles at end points, but in actual display they are drawn in magenta, so that it is easy to distinguish.

(2) Warnings for missing elements

The number of the missing elements (the elements in the model view that are not found in the student's view) is reported for each views in the model drawing, in the lower left corner of the student's drawing.

(3) Warnings for unmatched views

If the highest score for a student's view divided by the number of element is lower than the "allowable error ratio" mentioned in section 3.4, the bounding-box of the view is displayed in yellow lines. And the name of the model views that are not found in student's drawing are reported in the lower left corner of the drawing.

We also prepare two levels of comparison, "normal" and "strict". In the normal comparison, the elements of same geometry are regarded as the same. For example, arcs on the same circle are regarded as the same, regardless of the positions of the end points.

On the other hand, in the strict mode, the elements with different end points are judged as different, except for center lines. This is useful for detecting line segments of wrong length, but has a significant drawback. In figure 6, an example result of normal comparison, a fillet of wrong rounding size is detected and warned.

But with the strict mode, the adjacent vertical and horizontal line segments will also be warned, which will make it difficult for students to find out the fault.

We advise students to first use the normal mode, and then to try the strict comparison.

4. APPLICATION AND EVALUATION

We have been using this checking system in the introductory courses for mechanical engineering drawing at the Department of Mechanical Engineering, Tokyo Denki University since 2009. By introducing the checking system, the quality of students' drawings improved considerably. For example, the mistake described in figure 2 and 6 found in students' final drawings in 2009 was only 8 out of 126 drawings (6.3%), where it was 45 out of 108 (41.8%) in 2008.

This difference is partly because the error is rather subtle and teachers sometimes overlook it. But the main reason seems to be, with the automated checking system, that students' drawings have better chance to be checked for two reasons. Firstly, as students can check their own drawings for themselves, the checking can be done at any time, not only in the class. Secondly, students can complete the drawings faster, because they can check their drawings' geometry view by view, reducing the reworks. There are less incomplete drawings in 2009 than in 2008.

As for reducing waiting time for checking by teachers, significant change has not been observed. Unfortunately, there still appears a long queue of students waiting for their turn. This seems to be because checking capacity is in a saturated situation. (Four to five instructors are committed to the checking of 60-70 students per class.) There are many check points other than the object geometry, and most of the students want to be checked at a time, and for several times.

5. CONCLUSION AND FUTURE WORK

In order to improve the quality of checking and to reduce the waiting time for checking in-

introductory courses for technical drawings, we have developed an automated checking system which runs on CAD system. This system enables students check their drawings for themselves, regardless of the viewport layout.

Introducing this system to the introductory mechanical drawing courses in Tokyo Denki University showed that the system is effective for improving the quality of students' drawings. It also reduces instruction time for students, but did not reduce the waiting time for checking.

As future works, we plan to improve the system so that it can check center lines of improper length, because many students in the introductory courses tend to leave very long center lines. We also plan an automated checking of dimensions and other annotations, because this is the main concern in the current checking by teachers.

ACKNOWLEDGMENTS

I appreciate my colleagues who collaborate with the technical drawing courses at the Department of Mechanical Engineering, Tokyo Denki University, for giving a chance to introduce and to evaluate this checking system.

REFERENCES

- [1] N. Yamamoto, A. Shimura, S. Kumagai and Y. Honjo. A check method in the personal computer based CAD. *Journal of Japan Society for Design Engineering*, **35** (5): 178-184, 2000. (Japanese)
- [2] Y. Nagasaka, N. Katoh and H. Ohtaki. Development of CAD Drawing Check system for Education. *Journal of Japan Society for Design Engineering*, **36** (5): 221-228, 2001. (Japanese)

ABOUT THE AUTHOR

Ichiro TANAKA

Department of Mechanical Engineering,
Tokyo Denki University, Japan
tanaka@cck.dendai.ac.jp

SELF-REPEATING TRAJECTORIES OF LIGHT RAYS

Michael MANEVICH, Elizabeth ITSKOVICH and Nir SHVALB

Ariel University, Science Park, 40700, Israel

ABSTRACT: We consider the problem of constructing self-repeating trajectories of light rays inside a planar closed contour from the view point of projective geometry. Because we consider the trajectory of a light ray as a straight line, this problem on a plane is analogous to the billiard ball motion problem. Using the example of a trapeze we show that the projective correspondences of the rows of points and double points of these correspondences give us the opportunity to build only a closed contour, as some closed loop of endless trajectory. To construct the self-repeating trajectories, consisting of three, five or more segments, we suggest a new method. This method is based firstly, on the straightening a trajectory by means of reversing the whole construction and secondly, on the creating of two projective pencils of lines. The first pencil of rays corresponds to a set of rays that emanate from given points on the some line m of a contour under certain angles. The second pencil of rays corresponds to a set of rays returning to the same positions with the same angles. Self-repeating trajectory corresponds to the line that belongs to both pencils. This line passes through the double point of two projective rows of points on the line m . The suggested method for constructing the self-repeating trajectories is naturally generalized to the three-dimension space. In this case the collineation of two sets of points is set on the plane. The double points of this collineation indicate the existence of the self-repeating lines.

Keywords: Self-repeating trajectories, Projective correspondences, Double points.

1. INTRODUCTION

For beginning, we recall some concepts, definitions and provide some constructions which are known in the literature, for example: [4], [7] and [13].

- 1.1. Perspectivity and projectivity are marked respectively by symbols $\overline{\wedge}$ and $\overline{\vee}$.
- 1.2. Forms of the first grade:
 - 1.2.1. Rectilinear row of points, i.e. set of points, belonging to the same straight line.
 - 1.2.2. Pencil of straight lines, i.e. set of straight lines on a plane passing through a given point.
 - 1.2.3. Pencil of planes, i.e. set of planes passing through a given straight line.
- 1.3. The affine correspondence of planes is the correspondence formed by the chain of successive parallel projections. The invar-

iants of this correspondence are as follows:

- 1.3.1. Collinearity is when a straight line of one plane corresponds to a straight line of the other plane.
- 1.3.2. Mutual belonging of a point and a straight line.

Simple ratio of three points on the line:
 $(ABC) = AC/BC$

1.4. Collineation

A collineation of two fields is a one-to-one correspondence of their elements, in which the following is true:

- 1.4.1. Every point of one field has a corresponding point in the other field.
- 1.4.2. Every straight line of one field has a corresponding straight line in the other field.
- 1.4.3. A point lying on a straight line in one field corresponds to a point lying on the corresponding straight line in the other

- field.
- 1.4.4. A collineation on a plane is called a homology if a corresponding pair of triangles satisfies Desargues theorem. Homology has a center and axis of homology.
 - 1.5. Perspective and projective rows of points. Perspective and projective pencils of straight lines.
 - 1.5.1. A row of points perspective to a pencil of straight lines is a section of this pencil by a straight line.
 - 1.5.2. A pencil is perspective to a row of points is obtained when the row is projected by the pencil.
 - 1.5.3. Two pencils are perspective if their corresponding rays intersect at points lying on the same line.
 - 1.5.4. Two forms of the first grade are projective if the cross ratios of any four pairs (A, B, C, D) of the first row v and (A^1, B^1, C^1, D^1) of the second row v^1 are equal: $(ABCD) = (A^1B^1C^1D^1) = (ABC)/(ABD) = (A^1B^1C^1)/(A^1B^1D^1)$.
 - 1.6. One of the definitions of a pencil of straight lines of the second order – straight lines which connect corresponding points of two projective rows.
 - 1.7. The ordered correspondence: if point M describes a row in some direction, then the corresponding point M_1 also describes a row in a certain direction. Equally directed projective rows with a common carrier (on one line) can have two double points, one double point or no double points. Opposite directed rows always have two double points.
 - 1.8. One of the definitions of a quadratic curve - the locus of intersection points of corresponding rays of two projective pencils.
 - 1.9. One of the definitions of a pencil of the second degree - the locus of lines joining corresponding points of two projective rows.
 - 1.10. The rays of a pencil of the second de-

gree intersect any two rays of this pencil in two projective rows of points.

- 1.11. Reflected rays. The well-known law of reflection:

1.11.1. Incident and reflected rays lie in the same plane as the normal to the reflecting surface.

1.11.2. The ray of light l_1 incident on the transparent surface Σ is reflected from it at the same angle θ (in absolute value). Angle θ_1 is counted off the normal N (Fig.1).

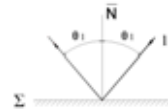


Figure 1: Law of Reflection.

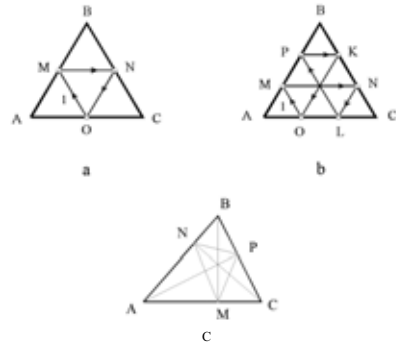


Figure 2: Self-Repeating Lines (periodic trajectory).

In Fig.2 a, b - light rays are reflected in an equilateral triangle. These figures were described earlier in [10].

In Fig.2a point O is the middle of segment AC . Ray l exits from this point at an angle of 60° to AC . Trajectory of this ray is the "self-repeating" triangle OMN . We have used the words "self-repeating" because the reflected ray, having come to point O , repeats the same trajectory again. In Fig.2b point O divides segment AC at the ratio of $AO/AC=1/3$

In geometry, Fagnano's problem is an optimization problem that was first stated by Giovanni Fagnano in 1775: For a given acute triangle determine the inscribed triangle of minimal perimeter. The orthic triangle MNP has the smallest perimeter of all triangles inscribed into an acute triangle [2] and [4]. "The orthic triangle MNP is a periodic trajectory since its angles are bisected by the altitudes of the triangle in which it is inscribed" [1].

It is not hard to notice that figure KLMN is parallelogram. Imagine, that pencil of light rays l_i^j exits from point K, where $i = 1, 2, 3, \dots, n$. For these rays we will get parallelograms with vertices R, T, F etc. using the same constructions. The question arises: what is the geometrical locus of these vertices? It is not difficult to answer this question. Firstly let us mark some pencils of light rays which we see in Fig.3.

gree (1.8).

It's easy to see that common ray $O_1 O_2$ corresponds to itself. In other words, if we take ray 1 as KA, then a corresponding ray in pencil O_1 will be $O_1 A$. Ray 1₄ will be coincident with KB and corresponding ray in pencil O_2 will be $O_2 B$, that is the same ray $O_1 O_2$ (in this case ray is considered as a straight line). Consequently, curve m^2 splits up into two straight lines: m and $O_1 O_2$.

Straight line m goes through points M, R, T, F, etc. (Fig. 3). Straight line m intersects side CD in point Q.

It's not hard to see from drawing in Fig.3, that segments $O_1 F = O_2 F$, $O_1 Q = O_2 Q$, $O_1 R = O_2 R$, $O_1 M = O_2 M$. As a result of equality of triangles, for example $\triangle O_1 QR = \triangle O_2 QR$ follows, that straight line m is bisector of angle $O_1 Q O_2$. In point Q angles of incident and reflected rays are equal as required.

Lastly, trajectory KPQSK is a self-repeating trajectory inside polygon ABCD.

Let us look closely at point Q. From the point of view of projective geometry this point is a double point of two projective rows of points on straight line CD, which are formed as the intersection CD with rays of two projective pencils.

We came to the classical problem of projective geometry: build intersection points of a straight line with a curve of the second degree, given by two projective pencils.

3. A CLOSED CONTOUR OF LIGHT RAYS IN SOME TRAPEZE

Let's consider trapeze ABCD (Fig. 4). Let us construct point O symmetric to point K relatively side CD. And similarly let's construct point P symmetric to point K in relation to side AB. Suppose that pencil of rays exits from point K and these rays intersect side AB. Let's construct light pencil with center P symmetrical to pencil of rays with center K regarding AB.

All further constructions of perspective and projective pencils similar to the previous constructions are considered in section 2. As a result of these constructions we have two projective pencils with centers O and P.

Let's construct the intersection points, as M, G, F and others of the corresponding rays of these projective pencils. The locus of these points is a curve of the second degree m^2 . The graphical construction of curve m^2 on separate points is an approximation. Consequently, intersection points of curve m^2 with a straight line AD gives us only an approximate answer.

It turns out that the intersection points mentioned above are the two double points of the two projective rows of points on the line AD. These rows are constructed by two projective pencils with centers O and P. Therefore, we can build these points exactly without using curve m^2 . One of these points is point Q which is shown in Fig.4.

This again emphasizes the beauty and elegance of projective geometry in solving both theoretical and practical problems.

Construction of double points of projective correspondence is known in the literature, such as [4] and shown in Fig. 4. We have a closed loop inside the trapeze. Let's explain:

Let's assume, that n – normal at point Q to side AD and a light ray exits from this point towards point B at an angle α to the normal n . Being reflected sequentially from sides AB, BC, CD of the trapeze this ray returns to the same point Q, the angle β is not equal to the angle α .

This means that in the future this ray will describe a different trajectory. We only have a closed branch (loop) of some trajectory. Below we consider constructing self-repeating trajectory.

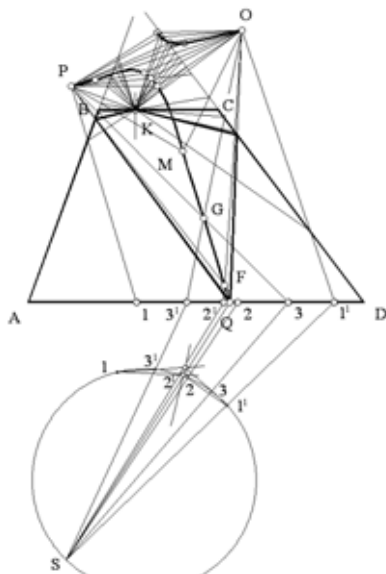


Figure 4: Closed Loop Inside the Trapezoid

4. SELF REPEATING TRAJECTORIES IN SOME TRIANGLE

Let's consider triangle ABC (Fig.5a).

We will now construct self-repeating trajectory NRP as shown in Fig. 2c.

The process of constructing this trajectory will be based on the projective correspondences rows of points and some new special designs. A light ray is reflected from sides AB, AC and BC of the triangle. Rays NR, RP and PN are created respectively. Let's construct a series of reflections of triangle ABC, in relation to sides AB, AC and BC. We see that the light trajectory is a straight line. The idea of such a "rectification" of a trajectory is not new and dates back to the theory of the motion of a billiard ball as described in the works, for example [1], [8], [11].

Let's take on side AB the series of points 1, 2, 3.... As a result, of reflections we obtain points $1_3, 2_3, 3_3, \dots$ on side of A_3B_3 . Since the distance is preserved, the rows of points projective (1.5.4).

These rows of points are in projective correspondence. Connecting the corresponding points by straight lines $1, 1_3; 2, 2_3; 3, 3_3$ we get pencil ω of the second order (1.9). Consider, for example, straight line of pencil ω , passing through points: 1 and 1_3 . This line "came" from point 1 on line AB at an obtuse angle to side $1A$ (angle $A11_3$) and returned to the same point 1_3 at an acute angle to side B_31_3 (angle B_31_31). If the line $1, 1_3$ is a self-repeating line the angles must be equal. Therefore construct beam of lines ω_3 passing through points $1_3, 2_3, 3_3, \dots$, so, that the angles of inclination to side A_3B_3 equal the angles, shown in Fig.5a. In other words, we overturn (or reverse) the entire structure, as shown in (Fig. 5a). It is easy to see that pencil ω_3 is congruent to pencil ω and row of points $1_4, 2_4, 3_4, \dots$ congruent row of points $1_3, 2_3, 3_3, \dots$. The points $1_4, 2_4, 3_4, \dots$ lie on line m_4 and on this line lies side AB of the triangle after the last reflection (Fig.5b, on this figure indexes are not specified so as not to overload the drawing).

Lines of pencil ω_4 connecting corresponding points of rows 1, 2, 3, and $1_4, 2_4, 3_4, \dots$ are parallel. Let's explain: Let's assume that point 1 is located on m_3 . The angle of ray $1-1_3$ with line m is the angle between lines m and m_3 . Then the ray from pencil ω_3 is m_4 should pass at the same angle to line m_3 . The challenge is to find a straight line belonging to two pencils ω and ω_3 . Then the desired line will belong to ω_4 (Fig.5a). The following chain projective correspondences follow from constructions in (Fig. 5a), (1.5.4) is follow:

$$\begin{aligned} & \overline{m(1,2,3,\dots)} \wedge \overline{m_3(1_3,2_3,3_3,\dots)} \wedge \overline{m_4(1_4,2_4,3_4,\dots)} \\ & \wedge \overline{m_3(1_5,2_5,3_5,\dots)} \end{aligned} \quad (3)$$

Consequently, the rows of points on the line m_3 are projective:

$$\overline{m_3(1_3,2_3,3_3,\dots)} \wedge \overline{m_3(1_5,2_5,3_5,\dots)} \quad (4)$$

Double point N of these projective rows gives us a solution. The line passing through point N and parallel to straight lines of pencil ω_4 , defines self-repeating line NRP.

If the above double point does not get inside segment A_3B_3 , then there is no solution. This means that segment NN does not lie completely inside the area, limited by broken line $A-B-C-A_3-B_3-A$. The question of the existence of such a line remains open and requires further investigation.

The forward and reverse paths of the ray are indicated by arrows inside the triangle and the straightened trajectory (Fig. 5b).

5. SELF REPEATING TRAJECTORIES IN THE THREE DIMENSIONAL CASE

Building a self-repeating line for the three-dimensional case is based on the same principle of straightening trajectory.

But in this case, we have a spatial configuration, for example, instead of triangle, a tetrahedron can be considered. Instead of flat pencils of straight lines we consider spatial bunches of straight lines. Instead of congruent rows of points on lines we consider flat congruent fields of points on planes. As a result of the constructions similar to the constructions above, we obtain a collineation (1.4) of two fields, double points of which indicate the existence of self-repeating lines. Note that in the collineation (non-homologous, 1.4.4) there can be no more than three double points. Every collineation has at least one double point [5]. This means that there may be at least one solution, i.e. at least one spatial a self-repeating trajectory.

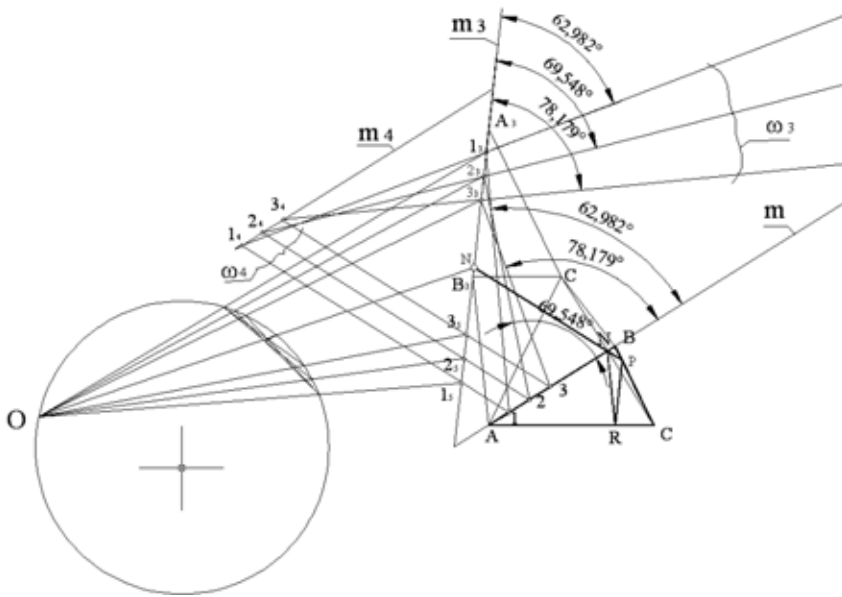


Figure 5a: Self Repeating Contour of Light Rays in the Triangle.

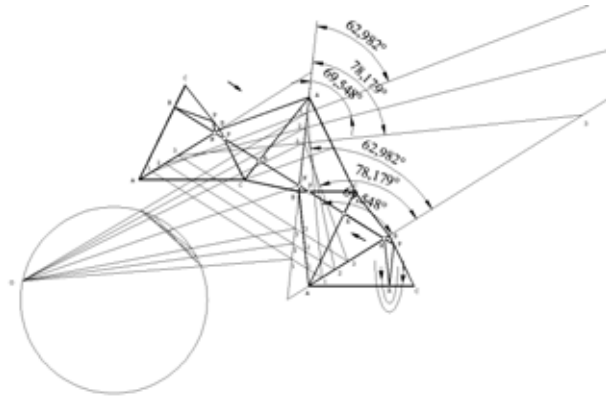
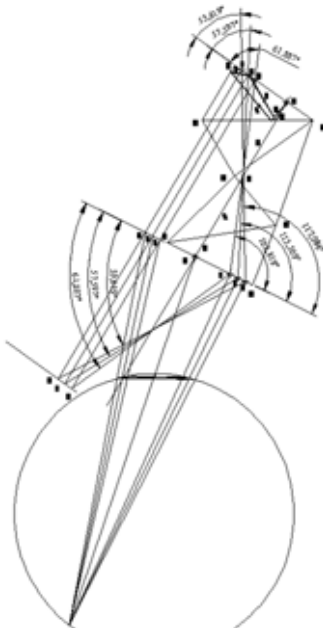


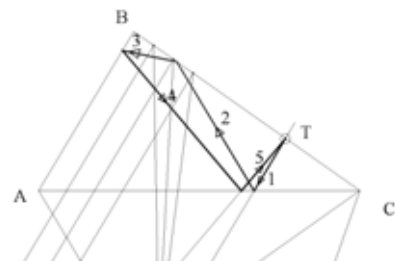
Figure 5b: The Forward and Reverse Paths of the Ray Inside the Triangle.

The construction of self-repeating trajectory comprising five segments 12345 inside triangle ABC similar to the previous constructions and

is shown in Fig. 6a. The trajectory 12345 in enlarged scale is shown in Fig. 6b.



a)



b)

Figure 6: Self- Repeating Contour with Five Segments in the Triangle.

6. CONCLUSIONS

We suggest a method for construction of closed and self-repeating trajectories in a planar polygon. We have proved and given examples that such trajectories are induced by double points of projective correspondences of two rows of points. These points must be within one of the polygon's edges. If we consider a general case, when a polygon is multiply reflected, the existence of a double point within one of polygon's edges is not sufficient for existence of

self-repeating trajectory. In order to receive such a trajectory the resulting ray (trajectory deployed in a straight line) should be confined within the unfolded polygon or in other words, within the area covered by all the reflected polygons. Although the solution of the last question is hard to attain, the authors believe that a partial answer for the question of existence of self-repeating trajectories is within our reach and should be considered in future.

REFERENCES

- [1] Andrew M. Baxter and Ron Umble. Quotation from Periodic Orbits of Billiards on an Equilateral Triangle. arXiv: math/0509292v7 [math.DS] 1. 2007.
- [2] A. Bogomolny, "Fagnano's Problem: What is It?" <http://www.cut-the-knot.org/Curriculum/Geometry/Fagnano.shtml>
- [3] B. Ben-Moshe, M. Shaniand ,N. Shvalb , "Efficient model for indoor radio paths computation", Simulation Modelling Practice and Theory, 29, (2012), pp. 163172.
- [4] A. Bogomolny, Fagnano's Problem: Schwarz's solution from Interactive Mathematics Miscellany and Puzzles <http://www.cut-the-knot.org/Curriculum/Geometry/Fagnano.shtml>, Accessed 06 June 2014.
- [5] N. Chetverukhin. Projective geometry, (in Russian). Prosvescheniye Publ.Moscow, pages 105-136, 189-191, 1969.
- [6] T.E. Faulkner, "Projective Geometry", TextitOliver and Boyd , Edinburgh and London, 1960.
- [7] N. Glagolev, Projective geometry. Vischay Shcola, Moscow, pages 59-75, 1963.
- [8] E. Gutkin, "Billiards in polygons", Physica D: Nonlinear Phenomena, Volume 19, Issue 3, April 1986, Pages 311–333
- [9] F. Holt "Periodic reflecting paths in right triangles.", Geom. Dedicata,46, pages 73-90, 1993.
- [10] M. Manevich, E. Itskovich. Light ray trajectories and projective correspondences. Journal for Geometry and Graphics. Volume 15, Number 2, pages 181-193, 2011.
- [11] T. Ruijgrok "Periodic orbits in triangular billiards.", Acta Phys. Polon., 22 ,pages 955-981, 1991.
- [12] R. Schwartz, "The Poncelet grid", Preprint, <http://www.springerlink.com/content/v700326231863619/>
- [13] O. Veblen, J.W. Young "Projective Geometry", Ginn and Co., Boston, 1910.

ABOUT THE AUTHORS

1. Michael Manevich Dept. of Industrial Engineering, Ariel University, Science Park, 40700, Israel, E-mail address: michael.manevich@gmail.com
2. Elizabeth Itskovich Dept. of Computer Science, Ariel University, Science Park, 40700, Israel, E-mail address: itskovichli@gmail.com
3. Nir Shvalb Dept. of Industrial Engineering, Ariel University, Science Park, 40700, Israel, E-mail address: nirsh@ariel.ac.il

SEVERAL SOURCES OF SHAPES

Szymon FILIPOWSKI

Cracow University of Technology, Poland

ABSTRACT: This paper presents an approach to realistic and effective interpretation and depiction of the real world with static images and shapes, in particular a scale of human perception and impact. The main assumption that “human natural interpretation and description of the world could be a chosen part of truth or only a result”, dictates the way of the next considerations. There is a scientific work based on searching for other ways of understanding the phenomenon of space starting from probability of symmetry as a method to investigate an effect of unknown factors. It has been proven that probability of symmetry is strictly connected with distance. Then there is an attempt to find more relations between different sets of a space.

Keywords: Shape, source, probability, symmetry, spatial factors

1. INTRODUCTION

The most popular methods of describing shape are those based on the regular division of static space into a grid. Generally a computer writes images as a matrix or functions of regular spaces. It affects the use of raster images consisting of pixels or vector files with basic functions describing points, curves and surfaces but limited by precision of location. Compression of data can be performed on text data or spatial relations. The method of vector description is analogous to traditional drawing in classic geometry. Those methods represent only the approximate effect of true objects or precision projections of unique objects. In classical geometry objects came from other geometrical objects or parts of Euclidean space but there are some examples like a point or a set with unknown origin. Geometry is commonly based on unexplained idea. The question has been raised, whether the real world is geometric. Shapes that people see as defined could be only a similarity of simple shapes. Moreover the position of elementary molecules which form matter is impossible to define. It means that the position and shape of an object in both big and small scale is only an approximation. It is unknown what the reason and effect of written re-

lations is. It is not obvious if a real space is an empty place filled with matter, fields and interactions or created by each of them. Our image of the world is controlled by intuition telling us that locally it is a 3-dimensional, static and metric space. Metric space means that there is a relation between every element of a set that could be measured. Human point of view dictates ordered interpretation of a space and organizes unpredicted information by rules of probability which are reduced to an understandable formula. Looking around we can understand only a small part of the entire world. The Euclidean space gives us a clear point of view. In effect we reconstruct in our minds predicted and unpredicted elements in a simplified space. This indicates the existence of something most general so that the author proposes another approach to the subject; setting a simple unique space only as an element of a more complex space. Above indicates that human natural interpretation and description of the world could be a chosen part of truth or only a result. While sketching a geometry construction or looking on reality we accept an effect of observation. There is an attempt to understand what are the reasons of this interpretation of space and to find true definition of those factors.

2. OBSERVATIONS

Some simple experiments were performed in search of general principles of space. A set was simulated by randomly spilling flour, wrinkling a sheet of paper, pouring hot wax into water and spreading ink in water. It created undetermined by human imagination natural forms placed on a plane and on 3D space. The 2d example of paper is not flat in fact but considering it from a topological point of view, we observe it being homogeneous to the plane but having a natural pattern. There are many similarities created through examples to each other and the real world but they were chosen to test only an example with paper because of continuity and analogy to the most popular method on thinking on a plane. All samples appear to be binary but analysis proves other properties.



Figure 1: Wrinkled paper as an example of general space.

A plain sheet of paper was closely creased and later unfolded. This was a way of achieving a random shape without the consideration of physical reasons. Looking closer on the achieved surface we can find an effect similar to a plane – an ideal situation.

This example shows continuous, metric space because it was created from that one. At first glance it looks chaotic but we can find some rules and connections between objects of the same size in a distance depending on size.

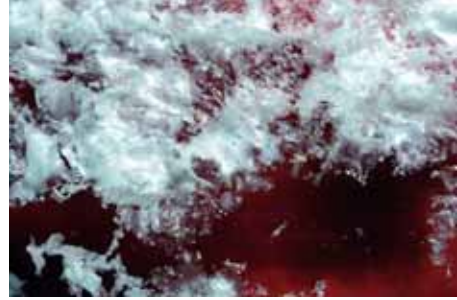


Figure 2: Random form of wax in a water- another example of free shape.

3. INTERPRETATION AND DESCRIPTION

In the probability of symmetry on continuous space where the less the distance is from the point of symmetry the less probability of the same variation.

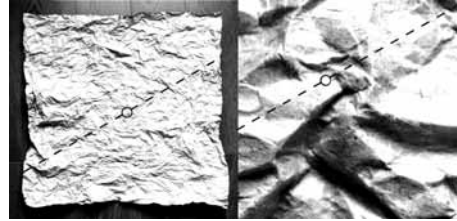


Figure 3: An area with investigated section and scope of the center.

The probability of symmetry means how probably is that one kind of property is symmetrical i.e. angle of plane parts and fracture. It could be described by the equation:

$$Q = \frac{1}{1 + (pd)^{2q}} \quad (1)$$

where Q is the probability of symmetry, d is a variable of distance which is a part of the full investigated interval and $0 \leq d \leq 1$; p , q are parameters steering the minimum value and progress of Q function, q is a positive integer. This method

of shape prediction should be supplemented by proceedings on edges of space or function by increasing the number of probes in a region where those changes are observed. A set of probability functions or probable variations is a metric space itself because we can measure the impact between probabilities from function Q used in the formula and all probable variations are connected. Above is a simulation of probability by a simply function.

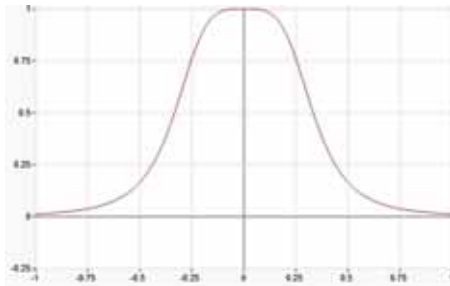


Figure 4: A graph of probability (1), where parameter $p = 3, q = 2$

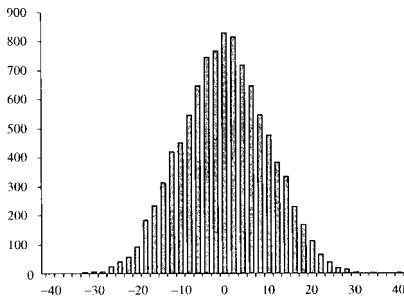


Figure 5: Gaussian distribution – scores of 6 dices. [6]

This is not a true shape but only an approximated interpretation which could also be described by exponential functions. This phenomena could be accurate described by Gaussian distribution, useful not only for statistic model of

solid objects. Quoting from a literature we can simulate a generalization of Brown movements as a way to understand basic kinematics analogous to throwing a dices. the static where the horizontal axis is a distance. The point of the matter is that the assemblage of scores is a Gaussian curve, but the effect in a space looks to be unordered.

The formula describes a situation:

$$D = \frac{1}{A} \times \frac{\sqrt{12}}{n} \times (Y1 + Y2 + \dots + Yn) - \sqrt{3n} \quad (2)$$

Where: A is a maximum of a score sum, n is the number of dices $Y1, \dots, Yn$ is a score of one throwing. With the assumption of big A, n .

Gaussian distribution could be adopted to describe different spaces like creased paper by defining a minimum limit of probability depending on a distance which is strictly connected with probability.

The question has been raised: what spatial transformation should be performed to modify a typical infinity curve like graph of shape probability to the limited space of Gaussian distribution?

First of all a method of thinking should be considered. Algebraic description uses the one property of distance which is written using a regular grid. This practical method is very simple and useful for analyzing only a section of a space or limited factors, but taking into the consideration more complex space relations only by one factor is complicated and sometimes could be impossible where more properties are encoded in this method of description.

Therefore a direct analysis of spatial factors has been subjected to reflection. The factors have also been compared.

Firstly a convexity impact on a side of space was considered. As an example an arc was used on a plane. A 1-dimensional objects lying on 2-dimensional objects divide into two parts. This rule could be applied to all dimensional objects or spaces. Where a less-dimensional object is finite it divides locally more-dimensional object into two sides. Where both are infinite or closed a rule of division exists on every scale. An arc not

ending on the investigated area was presented. It divides a plane into two sides. Taking to consideration other properties – a convexity we observe a strict impact of one property on another. On the top there is a positive convexity on the bottom a negative one so that a convexity here is dependant on sides and a side with convexity is dependant on distance.

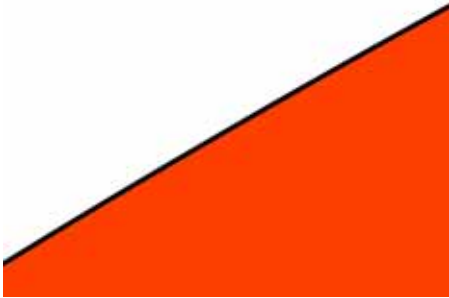


Figure 6: Example of a simple shape with sides, convexity and symmetry

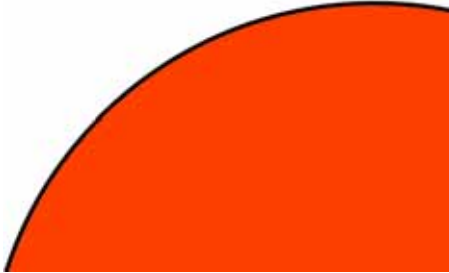


Figure 7: A bigger range of the same shape

Between them: scale of investigated area, accuracy of probing a space, and dimension. The question has been raised if a dimension is an equivalent factor like distance, convexity.

4. SHAPE IN FRACTAL GEOMETRY

Similar searching was performed on a fractal object – Koch curve. A character of an object deter-

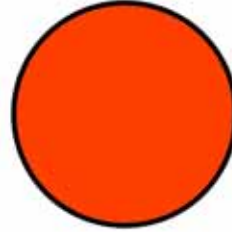


Figure 8: This scope shows relations between convexity, sides, symmetry, scale.



Figure 9: Changing a dimension (infinitive transformation).

mined a specific method of investigation based on the main assumption of fractal geometry – a self-similarity. It is an obstacle to define convexity using the classical method to fractal objects. It is required to investigate objects with precision appropriate for a scale of investigation. Calculations of convexity looking from one side gives same ratio on every steps of creating a form. It means that scope and precision becomes an important factor of a space interpretation.

Fractal geometry uses effect of a space, an object is analyzed as a new space where a point of view is situated on the original space. Point of view decides on some properties of a space it could be distinguished by dimension size and

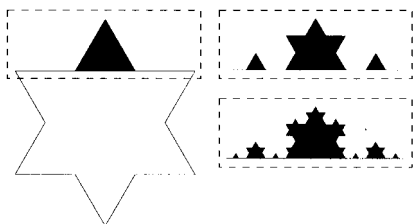


Figure 10: Convexity of a fractal structure: Koch Curve. Score: +2 and -1 of angles measuring from bottom for first step (left-top), +10 and -5 in 2nd (right-top), +42 and -21 in 3rd. [6]

position. Is it possible to investigate an object from itself or without any point of view? Projection in classical geometry, measuring in fractal geometry is a spatial process.

5. CONCLUSION

The main aim of the work has not been achieved yet, but the investigation resulted in important conclusions for the way to understand a space. The effect of observations is dictated by point of view. Connections of factors can tell us more about a space rather than a lone factor.

REFERENCES

- [9] G. Maresch: *Spatial Ability The Phases of Spatial Ability Research*, Journal for Geometry and Graphics Volume 17, 2013.
- [9] E. Fink, D. Wood: *Planar Strong Visibility*, International Journal of Computational Geometry & Applications, 2003.
- [9] S. Schuierer, D. Wood: *Visibility in Semi-Convex Spaces*, Journal of Geometry 60, 1997.
- [9] G.J.E. Rawlings, D. Wood: *Orthoconvexity and Its Generalizations*, Computational Morphology, 1988.
- [9] T. F. Banchoff: *From Flatland to Hypergraphics: Interacting with Higher Dimen-*

sions, Interdisciplinary Science Reviews, 1990.

- [9] H.-O. Peitgen, H. Jurgens and D. Saupe: *Fractals for the Classroom. Part 1: Introduction to Fractals and Chaos*, Wydawnictwo Naukowe PWN SA, 445-502, Warszawa 2002.
- [9] K. Bartel: *Perspektywa malarska*, Vol. I, Vol. II, PWN Warszawa, 1955.
- [9] I. Stewart: *FROM HERE TO INFINITY A Guide to Today's Mathematics*, Prószyński Media Sp. Z.o.o., Warszawa, 2012.
- [9] M. Sroka-Bizoń, E. Terczyńska: *PERCEPTION OF VIEW- HOW TO DEVELOP SPATIAL IMAGINATION*, The Journal of Polish Society for Geometry and Engineering Graphics, Volume 25, 2013.

ABOUT THE AUTHOR

Szymon Filipowski is an architect working in Division of Geometry, Technical Drawing and Engineering Graphics. Interested in mathematics and physics, especially in theory of space and perception.

SIMULATING THE CONSTRUCTION PROCESS OF SONG-STYLE DOUGONG USING DIGITAL GRAPHICS

Shilun HAO, Adrian H. TAN, Fabian H. TAN, and Frank M. CROFT Jr.

The Ohio State University, United States

ABSTRACT: The dougong is a trait of Chinese architecture and specifically the construction engineering of its roofing system, which can be translated into ‘cap and block’ bracket system. In this research, the Song-style dougong recorded in Yingzao Fashi, the oldest currently preserved construction specifications, are systematically reconstructed by using 3-D digital graphical techniques. During the modeling process, step-by-step construction operations of more than 20 different kinds of dougong and the delicate workmanship of the components are presented in detail. Then, the categories of Song-style dougong are compiled in a knowledge-based intelligent dougong system (IDS). The IDS presents graphical simulations of the construction process of the Song-Style dougong, based on the queries prompted to the users. It allows users to select variables such as dimensions, the number of levels, and the type of components, among others to find the type of dougong expected. Once selection processes are completed, the IDS furnishes the users with a conclusion in both textual and graphical form, and the step-by-step assembly of the dougong is shown from various angles. Through the analyses and observations of Song-style dougong in 3-D graphical environment, the beauty and wisdom of ancient Chinese architecture can be viewed and recognized; this study therefore introduces an important element in ancient Chinese architecture and a new attempt to reconstruct it in digital graphics.

Keywords: Chinese architecture, Computer graphics, dougong, expert system, computer aided design, graphic simulation.

1. INTRODUCTION

The term *dougong*, used as both singular and plural, contains the terms *dou* and *gong*. A *dou* is an inverted cap holding a *gong*, a bow-like block, on top of it. Each *dou* has a small hole drilled in the bottom into which a matching pin protruding from a *gong* is mated. Skillfully made *dougong* joinery does not need fasteners such as nails and glue. In essence, a *dougong* configuration is constructed to support an overhang on top of it. The earliest known *dougong* systems can be found in ceramic models of funerary buildings, especially those from the Eastern Han Dynasty, ca. AD 20 – 220 (Fig. 1). The earliest written documents of dougong system were found in the Chinese building standards called *Yingzao Fashi* which were first introduced in AD 1103 in the Song

Dynasty, ca. AD 960 – 1279, and in the Chinese construction manual and specifications named *Gongcheng Zuofa* in AD 1734 which were compiled and enforced during the Qing Dynasty, ca. AD 1644 - 1912. The *dougong* recorded in Yingzao Fashi is called the Song-style *dougong* while the *dougong* recorded in Gongcheng Zuofa is called the Qing-style *dougong* by Chinese architects. Since its conception, this support system flourished not just in China, but spread to Japan, Korea, and South East Asia.



Figure 1: Corner dougong both from Eastern Han Dynasty, ca. AD 25 – 220, Henan Provincial Museum, Zhengzhou, China

The *dougong* is not a mere decoration in ancient Chinese architecture, but a multipurpose structural component with the following functions:

- i. The total weight of the roof can be transferred to the pillar by the *dougong*, and then to the foundation through the pillar. Here, the *dougong* serves as a link between the superstructure and substructure.
- ii. The *dougong* can extend the most outside eaves up to a certain distance from the pillar. The extended eaves in turn protect the pillar from humidity and rain.
- iii. Owing to the fact that the *dougong* is regularly made of wood, it can be carved and decorated, and be used to decorate the eaves of the roof.
- iv. Due to the often exclusive style of components from certain dynasties in ancient China, the *dougong* can be used to identify the period and kingdom during which it was employed for building construction.
- v. The mortise-and-tenon joints of the *dougong* increases the roof integrity, especially with respect to seismic forces. A *dougong* can be considered a non-rigid component, so it can effectively absorb the energy from an earthquake. The best case

here is the Yingxian Wooden Pagoda (ca. AD 1056, 67-m high), China's oldest and highest structure entirely made of timber, which was struck by more than forty earthquakes [4,8] but still stands to this day.

Much research has been made about the *dougong*, primarily focusing on its mechanical performance [1,6] and seismic analyses [2,5]. Li et al. [3] used 3-D modeling techniques to reconstruct a timber structure building in the Tang-Dynasty. Still, no study has been conducted to create an intelligent graphical model of step-by-step *dougong* construction as presented in this paper. Thus, the objective of our study is to create a prototype of an Intelligent Dougong System (IDS) as a learning tool to understand the erection procedure of the Song-style *dougong*. The names of components are especially deep-rooted in the Chinese culture, deeming them inaccurate should they be translated into English. Hence, in this paper, we will keep the names as they are.

2. SONG-STYLE DOUGONG

In Yingzao Fashi, the word “*puzuo*” is synonymous to “*dougong*,” describing *dougong* made prior to the Qing Dynasty. In Gongcheng Zuofa, the Chinese building specification published and enforced in the Qing Dynasty, the name of the *dougong* is expressed as “*ke*.” In the following text, in order to be consistent with the original nomenclature in Yingzao Fashi, *puzuo* will be used to describe the Song-style *dougong* like Bujian Puzuo, Zhutou Puzuo and Zhuanjiao Puzuo.

2.1. The Units and Cai

The units used in Yingzao Fashi are mainly *fen*° and *Song Chi*. For dimension purposes, one unit of Song Chi is equal to ten units of *Song Cun*, which is about 1.22 inches or 30.9 millimeter in the metric system. One unit of *fen*° is equal to 0.6 unit of *Song Cun*, which is approximately 0.73 inches or 18.54 millimeter in the metric system.

In Yingzao Fashi, the dimensions of Song-style *dougong* are classified into 8 different scales based upon the overall size of the building. Each scale in this grading system corresponds to a range of sizes the components made. The first grade is called the “*Yideng Cai*,” which refers to the biggest components while the eighth grade stands for the smallest in dimension. Despite grade differences, the components generally have the same styles and configurations. In each grade of the *Cai*, only the width and the height of the material size are specified, while the length depends on the specified component and application in construction practice [7], as shown in TABLE 1 (all units in *fen*°).

TABLE 1: Rules for selecting the *Cai* and its dimensions in *Yingzao Fashi*

Grade Name	Width	Height	Design Function
First-class <i>Cai</i>	10	15	Used for the palace with nine to eleven spans
Second-class <i>Cai</i>	9.17	13.75	Used for the palace with five to eight spans;
Third-class <i>Cai</i>	8.33	12.5	Using for the palace with three to five spans; Using for the hall with seven spans
Fourth-class <i>Cai</i>	8	12	Used for the palace with three spans; Used for the hall with five spans
Fifth-class <i>Cai</i>	7.33	11	Used for the palace with three small spans; Used for the hall with five large spans
Sixth-class <i>Cai</i>	6.67	10	Used for pavilions and small hall
Seventh-class <i>Cai</i>	5.83	8.75	Used for small

Seventh-class <i>Cai</i>			palaces, pavilions
Eighth-class <i>Cai</i>	5	7.5	Used for small pavilions, cessions

2.2. The *Tiao*

A *dougong* is extended span-wise (*Tiao*) with the center line of the cross section of the *Lu-dou* (the first or the base *dou*) as its benchmark. Such inward or outward extensions are called “*Nei Tiao*” and “*Wai Tiao*,” respectively. In particular, the *tiao* extended from *Hua-gong* are called *chao*. An illustration of the *tiao* of one type of *dougong*, Bujian Puzuo of five levels with one *chao* and one *xia ang* outside and two *chao* inside (FIGS. 2 and 3).

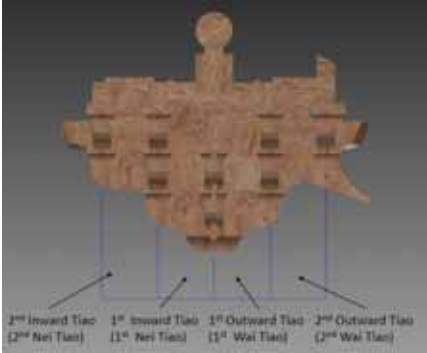


Figure 2: An illustration of *tiao* in Bujian Puzuo of five levels with one *chao* and one *xia ang* outside and two *chao* inside in front elevation.

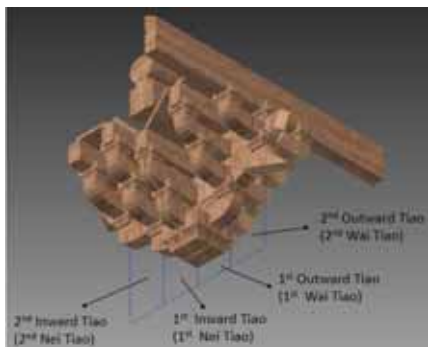


Figure 3: An illustration of *tiao* in Bujian Puzuo of five levels with one *chao* and one *xia ang* outside and two *chao* inside.

2.3. Nomenclature of Song-style *dougong*

In Song-style *dougong*, the names of the *dou* and *gong* in different positions have different names. This unique nomenclature system not only defines and specifies the dimension, position, function and appearance of *dou* and *gong*, but also implies profound Chinese traditions, serving as culture carrier in architecture.

An illustration of the introduction to nomenclature of Song-style *dougong* is shown in Figure 4, for the Bujian Puzuo of five levels with one *chao* and one *xia ang* in the outside and two *chao* inside as an example.

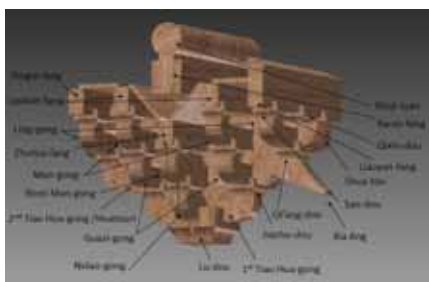


Figure 4: An illustration of components in Bu-jian Puzuo of five levels with one *chao* and one *xia ang* outside and two *chao* inside.

2.4. Compound components

In the Zhuanjiao Puzuo, there are two special types of compound components that combine the two basic components.

The first compound component is called the “A combined B” where A and B stand for two different components, each of which is exactly half the length of the original component, often with some minor decorative modifications. In addition, these components are always used in pairs and in a 45-degree (diagonal) direction within the Zhuanjiao Puzuo. In the following text, such a component will be written as “A / B,” for the limitation of the space. An example would be an illustration of *ling-gong* combined with *guazi-gong* used in the eight-level Zhuanjiao Puzuo with two *chao* and three *xia ang* (Fig. 5).



Figure 5: An illustration of combined components in *Zhuanjiao* Puzuo of eight levels with two *chao* and three *xia ang*.

The other type of compound component is called the *Yuan Yang Jiao Shou*, abbreviated as YYJS in the following text. The YYJS is the subset of the combined component stated above while the difference between the two is that there are two ends in the YYJS component connecting with another *puzuo* in the supporting system or with another complete basic component, as illustrated in Figure 6, a *man-gong* combined with the *qieji-tou* YYJS used in the eight-level Zhuanjiao Puzuo with two *chao* and three *xia ang*:

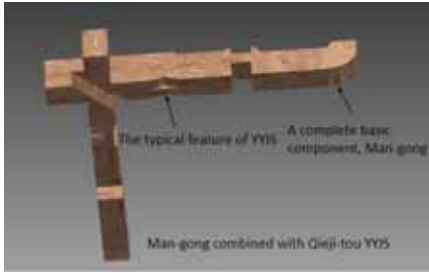


Figure 6: An illustration of *man-gong* combined with *qieji-tou YYJS* in *Zhuanjiao Puzuo* of eight levels with two *chao* and three *xia ang*.

3. CLASSIFICATION OF SONG-STYLE DOUGONG

In *Yingzao Fashi*, the classification of Song-style *dougong* is primarily based on the number of levels and functional positions. Dimension-wise, each type of *dougong* is classified into eight different grades based on its size. Based on calculations of the number of different types of *dougong* combination, there would theoretically be more than 140 different types of *dougong* recorded in the *Yingzao Fashi*.

The categories and types of Song-style *dougong* used for the IDS introduced in this paper follow those tabulated in Table 2.

TABLE 2: Categories and types of Song-style *dougong*

CA TE GO RY	CRITERIA	TYPE
A	GENERAL TYPE	<ol style="list-style-type: none"> 1. Dougong recorded in <i>Yingzao Fashi</i> 2. Dougong which existed in currently preserved architectures is not fully in accordance with the specifications in <i>Yingzao Fashi</i> 3. Simulated Song-style <i>dougong</i> used in modern architecture
B	NUMBER OF LEVELS	<ol style="list-style-type: none"> 1. Less than three levels 2. Four levels 3. Five levels 4. Six levels 5. Seven levels 6. Eight levels
C	DIMEN- SIONS	<ol style="list-style-type: none"> 1. Grade I (Largest) 2. Grade II 3. Grade III 4. Grade VI 5. Grade V 6. Grade VI 7. Grade VII 8. Grade VIII (Smallest)
D	SELECTED ANG	<ol style="list-style-type: none"> 1. Xia-ang (lower ang) 2. Shang-ang (upper ang)
E	POSITION OF <i>DOUONG</i>	<ol style="list-style-type: none"> 1. Bujian Puzuo 2. Zhutou Puzuo 3. Zhuanjiao Puzuo

4. CONSTRUCTION SIMULATION

The following is an example of the construction process of an eight-level Zhuanjiao Puzuo with three *chao* and three *xia ang*. The construction process will be presented level by level.

4.1. Level 1

Place the Jiao Lu-dou upon the pillar as the foundation or pedestal of the dougong (Fig. 7).

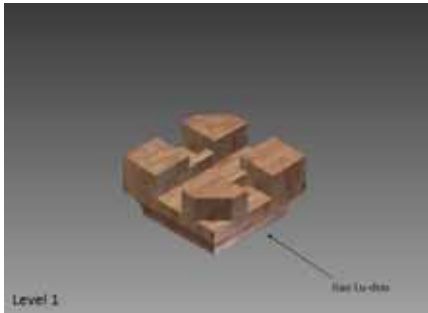


Figure 7: Setup of components in Level 1

4.2. Level 2

- Step 1: Place a pair of 1st *Tiao Jiao Hua-gong* / *Nidao-gong* into the *Jiao Lu-dou*.
- Step 2: Place the 1st *Tiao Jiao Hua-gong* into the 45-degree slot of the 1st *Tiao Jiao Hua-gong* / *Nidao-gong*.
- Step 3: Place two *Pingpan-dou* upon the ends 1st *Tiao Jiao Hua-gong*.
- Step 4: Place two *San-dou* upon the ends of the *Nidao-gong*, and two *Jiaohu-dou* on the ends of 1st *Tiao Hua-gong* (Fig. 8).

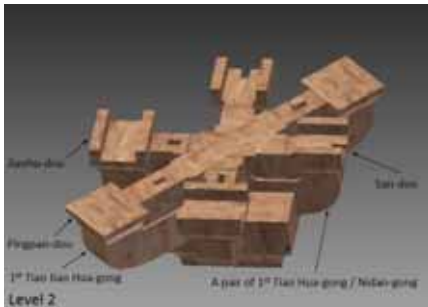


Figure 8: Setup of components in Level 2

4.3. Level 3

- Step 1: By inserting the mutually matched slots into each other, connect a pair of outer *Guazi-gong* / *Xiaogong-tou* with the 2nd *Tiao Jiao Hua-gong*.
- Step 2: Connect a pair of the 2nd *Tiao Jiao Hua-gong* / *Man-gong* with the 2nd *Tiao Jiao Hua-gong*.
- Step 3: Connect a pair of inner *Guazi-gong* / *Xiaogong-tou* with the 2nd *Tiao Jiao Hua-gong*.
- Step 4: Place the 2nd *Tiao Jiao Hua-gong* upon the 1st *Tiao Jiao Hua-gong*, and let the components of Level 3 match the *San-dou* and *Jiaohu-dou* on Level 2.
- Step 5: Place the *San-dou*, *Jiaohu-dou* and *Pingpan-dou* on their corresponding positions on Level 3 (Fig. 9).

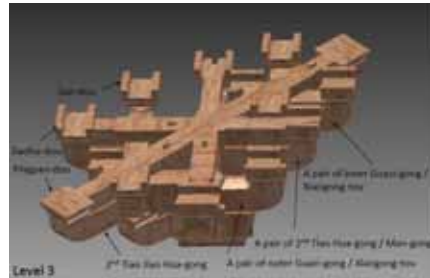


Figure 9: Setup of components in Level 3

4.4. Level 4

- Step 1: By inserting the mutually matched slots into each other, connect a pair of *Guazi-gong* / *Xiaogong-tou* “YYJS” with the 3rd *Tiao Jiao Hua-gong* / *Huatouzi*.
- Step 2: Connect a pair of *Man-gong* / *Qieji-tou* with the 3rd *Tiao Jiao Hua-gong* / *Huatouzi*.
- Step 3: Connect a pair of *Huatouzi* / lower *Zhutou-fang* with the 3rd *Tiao Jiao Hua-gong* / *Huatouzi*.
- Step 4: Connect a pair of *Man-gong* / *Qieji-tou* “YYJS” I with the 3rd *Tiao Jiao Hua-gong* combined with *Huatouzi*.

- Step 5: Connect a pair of *Guazi-gong* / *Xiaogong-tou* with the 3rd *Tiao Jiao Hua-gong* / *Huatouzi*.
- Step 6: Place the 3rd *Tiao Jiao Hua-gong* / *Huatouzi* upon the 2nd *Tiao Jiao Hua-gong* and let the components of Level 4 match the *San-dou* and *Jiaohu-dou* on Level 3.
- Step 7: Place the *San-dou*, *Jiaohu-dou* and *Pingpan-dou* on their corresponding positions on Level 4 (Fig. 10).

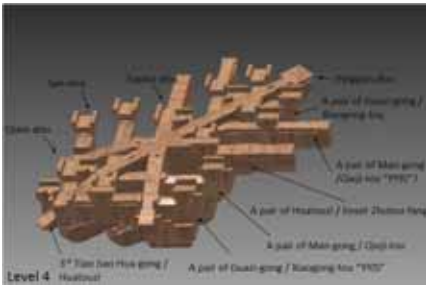


Figure 10: Setup of components in Level 4

4.5. Level 5:

- Step 1: By inserting the mutually matched slots into each other, connect a pair of *Guazi-gong* / *Xiaogong-tou* with the *Jiao Ang I* / *Jiao Shua-tou*.
- Step 2: Connect a pair of outer *Luohan-fang* with the *Jiao Ang I* / *Jiao Shua-tou*.
- Step 3: Connect a pair of the *Xia Ang I* / mid-lower *Zhutou-fang* with the *Jiao Ang I* / *Jiao Shua-tou*.
- Step 4: Connect a pair of inner *Luohan-fang* with the *Jiao Ang I* / *Jiao Shua-tou*.
- Step 5: Connect a pair of *Man-gong* / *Qieji-tou* "YYJS" with the *Jiao Ang I* / *Jiao Shua-tou*.
- Step 6: Connect a pair of *Ling-gong* / *Xiaogong-tou* with the *Jiao Ang I* / *Jiao Shua-tou*.
- Step 7: Place the *Jiao Ang I* / *Jiao Shua-tou* upon the 3rd *Tiao Jiao Hua-gong* on Level 4 and let the components of Level 5 match the *San-dou* and *Jiaohu-dou* on

Level 4.

- Step 8: Place the *San-dou*, *Qi'ang-dou* and *Pingpan-dou* on their corresponding positions on Level 5 (Fig. 11).

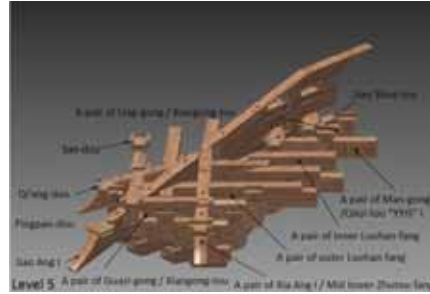


Figure 11: Setup of components in Level 5

4.6. Level 6

- Step 1: By inserting the mutually matched slots into each other, connect a pair of *Guazi-gong* / *Xiaogong-tou* with the *Jiao Ang II*.
- Step 2: Connect a pair of the *Xia Ang II* / Mid upper *Zhutou-fang* with the *Jiao Ang II*.
- Step 3: Place the two *Guazi-gong* upon the *Xia Ang II* on the right side and on the left side.
- Step 4: Connect a pair of inner *Luohan-fang* with the *Jiao Ang II*.
- Step 5: Connect a pair of *Man-gong* / *Qieji-tou* "YYJS" II with the *Jiao Ang II*.
- Step 6: Place the *Jiao Ang II* upon the *Jiao Ang I* on Level 5 and let the components of Level 6 match the *San-dou* and *Jiaohu-dou* on Level 5.
- Step 7: Place the *San-dou*, *Qi'ang-dou* and *Pingpan-dou* on their corresponding positions on Level 6 (Fig. 12).

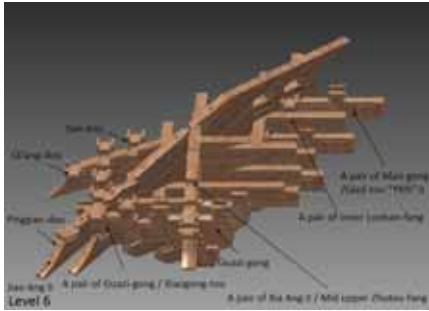


Figure 12: Setup of components in Level 6

4.7. Level 7

- Step 1: By inserting the mutually matched slots into each other, connect a pair of *Guazi-gong* / *Xiaogong-tou* with the *Jiao Ang III*.
- Step 2: Connect a pair of *Man-gong* / *Qieji-tou* “YYJS” III with the *Jiao Ang III*.
- Step 3: Connect a pair of outer *Luohan-fang* with the *Jiao Ang III*.
- Step 4: Connect a pair of the *Xia Ang II* / Upper *Zhutou-fang* with the *Jiao Ang III*.
- Step 5: Place the two *Guazi-gong* upon the *Xia Ang III* on the right side and on the left side.
- Step 6: Connect a pair of inner *Luohan-fang* with the *Jiao Ang III*.
- Step 7: Connect a pair of *Pingqi-fang* with the *Jiao Ang III*.
- Step 8: Place the *Jiao Ang III* upon the *Jiao Ang II* on Level 6 and let the components of Level 7 match the *San-dou* and *Jiaohu-dou* Level 6.
- Step 9: Place the *San-dou*, *Qi'ang-dou* and *Pingpan-dou* on their corresponding positions on Level 7 (Fig. 13).

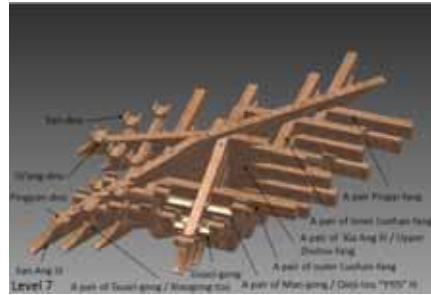


Figure 13: Setup of components in Level 7

4.8. Level 8

- Step 1: By inserting the mutually matched slots into each other, connect a pair of *Guazi-gong* / *Ling-gong* with the *You Ang*.
- Step 2: Connect a pair of *Man-gong* / *Qieji-tou* with the *You Ang*.
- Step 3: Place the two *Shua-tou* upon the *Xia Ang III* on the right side and on the left side.
- Step 4: Place the two *Ling-gong* connecting with *Shua-tou* and upon the *Xia Ang III* on the right side and on the left side.
- Step 5: Place the two *Man-gong* connecting with *Shua-tou* and upon the *Xia Ang III* on the right side and on the left side.
- Step 6: Connect a pair of outer *Luohan-fang* with the *You Ang*.
- Step 7: Place the *You Ang* upon the *Jiao Ang III* on Level 7 and let the components of Level 8 match the *San-dou* and *Jiaohu-dou* on Level 7.
- Step 8: Place the *San-dou*, *Qi'ang-dou*, *Qixin-dou* and *Pingpan-dou* on their corresponding positions on the Level 8 (Fig. 14).

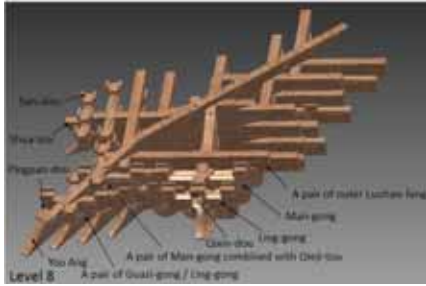


Figure 14: Setup of components in Level 8

4.9. Top

- Step 1: Place a pair of *Liaoyan-fang* upon *Gua-zi-gong* / *Ling-gong* and *Ling-gong* on the *Xia Ang III*.
- Step 2: Place a pair of *Luohan-fang* upon the *Man-gong* / *Qiej-tou* and *Man-gong* on the *Xia Ang III*.
- Step 3: Place a pair of *Chenfang-tou* upon *Shua-tou* on the right side and on the left side.
- Step 4: Place a *Baoping* on the headmost *Pingpan-dou* (Fig. 15).

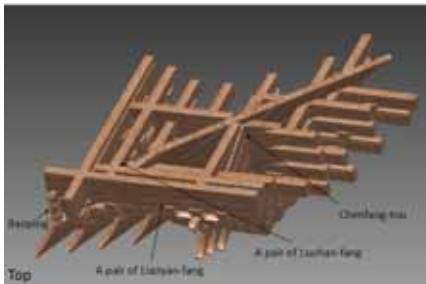


Figure 15: Setup of components in the top

The final results of the Song-style *dougong* are shown below in Figures 16, 17 and 18, viewed from top, front, and rear, respectively:



Figure 16: Top view of Song-style dougong

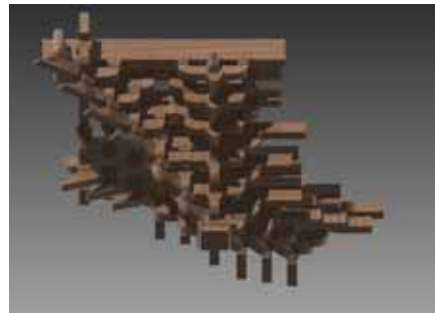


Figure 17: Front view of Song-style dougong

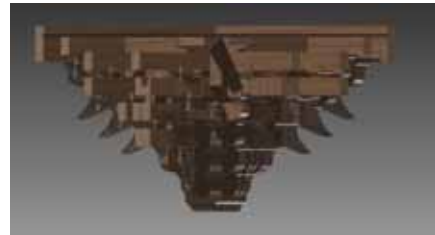


Figure 18: Rear view of Song-style dougong

5. INTELLIGENT DOUGONG SYSTEM (IDS)

After the modeling is complete, the categories of Song-style *dougong* are compiled in a knowledge-based intelligent *dougong* system (IDS). Not only does the IDS present all the simulation of construction process of Song-style *dougong*, but also, based on the queries prompted to the users, it will allow them to select the variables such as the dimensions, the number of levels, the type of *ang*, etc.,

to find the type of *dougong* expected.

The IDS includes a welcome screen (Fig. 19), an introduction section in which users can select introductive information of Song-style *dougong* and view the related illustrations (Fig. 20), along with an introduction of components (Fig. 21) in which users can view the definition of components, a multi-view drawing, and a 3D-graphics drawing shown from different angles. In the USER SELECTION section (Fig. 22), the user can select their desired Song-style *dougong* by answering the questions prompted.



Figure 19: The welcome screen



Figure 20: General introduction section



Figure 21: Introduction of components



Figure 22: User selection screen

Once the selection processes are completed, the IDS provides a conclusion related to the construction process, in both textual and graphical form. The step-by-step erection of dougong is then shown in both isometric and exploded isometric views from various angles (Fig. 23).



Figure 23: Construction process simulation

The IDS is a program which compiles all information of construction process of Song-style *dougong* in a manner akin to a knowledge base. Indeed, using analyses with the decision tree model, the IDS can help users find the desired *dougong*.

6. CONCLUSION

This research essentially finishes a knowledge-based of modeling of Song-style *dougong* recorded in the Yingzao Fashi. Based on this knowledge-base, an intelligent *dougong* system (IDS) was developed and introduced for the first time in this paper. When using the system, users can view the construction process, assembly method, component connections, and their configurations both graphically and textually. Also, based on the model of a decision tree, IDS can assist the users in their selection of

dougong they desired for a specific *dougong* category.

In addition, animations are included in the IDS to present the construction process of *dougong* dynamically and smoothly. In future studies, the authors will expand this research to the Qing-style *dougong* recorded in Gongcheng Zuofa, to complete the systematic 3D modeling work of the *dougong* in ancient Chinese architecture.

ACKNOWLEDGEMENTS

The first author expresses his gratitude and love to his parents who support him in this study and to whom this paper is dedicated. The second author dedicates this study to his parents, who encouraged his efforts and enabled him to provide the information needed to make this system possible.

REFERENCES

- [1] Chang, J. "Finite element modeling of Yingxian Wooden Tower." Beijing University of Technology, Master's Thesis, 2009.
- [2] Du, L., H.Li, F.Xue, and D.Qin. "The study of seismic behavior of Yingxian wooden pagoda," pages 363-370. China Civil Engineering Journal 43, 2010.
- [3] Li, D, T.Hong, T.Zhu, and J.Yang. "3D reconstruction and simulating assembly of ancient Chinese timber-structure building." CIPA XX International Symposium, 2005.
- [4] Li, H., F.Xue, and D.Qin. "The analysis and application of the dynamical properties of dougong in timber structure," pages 1655-1660. Industrial Building, 2006.
- [5] Sui, Y., H.Zhao, J.Xue, X.Zhang. "Experimental Study on Stiffness of Dougong in Chinese Ancient Buildings," pp. 819-822. Advanced Materials Research, 2012.
- [6] Zhang, T. and F.Lu. "The Dynamical Property Analysis of Ancient Wooden Pagoda," pages 81-86. Engineering Mechanics 21.1, 2004.
- [7] Zhou, Y. "An introduction to the characteristics of Song-style dougong," pages 55-56. Public Arts, 2010.
- [8] Zhu, E., Z.Chen, J.Pan, and L.Wang. "Finite Element Modeling of the Structural Performance of Dougong Brackets of Yingxian Wooden Pagoda." World Conference on Timber Engineering, 2010.

ABOUT THE AUTHORS

1. Shilun Hao is a graduate student at the Ohio State University. Shilun has a B.S. in Civil Engineering from the Beijing University of Civil Engineering and Architecture and is currently working towards an M.S. in Civil Engineering at the Ohio State University, with a focus on the digital reconstruction of ancient architecture.
2. Adrian H. Tan is a graduate student at the Ohio State University. Adrian has a B.S. in Computer Science and Engineering and an M.S. in Civil Engineering from the Ohio State University. Adrian is currently working towards a Ph.D. in civil engineering and construction with a focus on computer graphics and virtual simulation in the engineering industry.
3. Dr. Fabian H. Tan is a Professor in the Department of Civil and Environmental Engineering and Geodetic Science at The Ohio State University, where he has been since 1982. Dr. Tan has an M.S. in structural engineering, an M.E. in construction engineering and management, and a Ph.D. in construction engineering and management from the University of California in Berkley. Dr. Tan has served as a forensic consultant with numerous law firms in Ohio as well as the United States Air Force Weapons Laboratory in Kirtland, New Mexico. Additionally, Dr. Tan has also served as a project manager and engineer on multiple in-

ternational construction projects such as Leighton Australia-Indonesia Construction Inc. and Mahkota-Ekman Sweden Inc. Dr. Tan is currently teaching History of Ancient Engineering at The Ohio State University.

4. Dr. Frank M. Croft Jr., P.E., is an Associate Professor in the Department of Civil and Environmental Engineering and Geodetic Science at The Ohio State University, where he has been since 1984. Dr. Croft has a B.S. in Aerospace Engineering from the Indiana Institute of Technology, an M.S. in Engineering from the West Virginia College of Graduate Studies, and a Ph.D. in engineering at Clemson University in South Carolina. Dr. Croft has been very active with the Engineering Design Graphics Division (EDGD) of the American Society for Engineering Education (ASEE) since 1974.

SIMULATING THE CONSTRUCTION PROCESS OF THE ROMAN COLOSSEUM USING DIGITAL GRAPHICS

Adrian H. Tan, Frank M. Croft Jr., Fabian H. Tan
The Ohio State University, United States

ABSTRACT: This research focuses on a digital step-by-step recreation of the construction process of the Colosseum, the famous ancient Roman amphitheater, demonstrating that the process of retracing the construction of such a large and complex monument presents a variety of challenges. Computer-generated imagery, or CGI, has been used to recreate ancient structures based on literature and archaeological evidence, focusing primarily on completed structures during their period of use. Given enough data, computer graphics can serve as a tool in simulating the construction of ancient monuments as well, though this approach warrants a balance between optimization and accuracy. Provided with extensive literary research and on-site analysis, the simulation discussed uses general-purpose engineering graphics software which provides an ideal balance of user-friendliness and complexity handling. The creation of the model, using the functionality of the software in question, reveals significant potential as an educational tool for understanding the enigmatic construction processes of the Colosseum, as a training tool for the construction operations of ancient edifices, and as an investigative tool for renovating such structures. Still, setbacks such as complexity spikes and erroneous feature mapping, owing to the dissonance between the computation capabilities of the hardware, the functionality of the software, and the data values provided need to be overcome.

Keywords: Computer graphics, digital modeling, computer aided design, architecture, archaeology, Colosseum.

1. INTRODUCTION

With the advent of modern technology, the creation and analysis of architectural blueprints and simulations is becoming a second-hand task, particularly in light of advancement in Computer-Generated Imagery (CGI). One aspect of digital reconstruction involves the restoration of ancient structures and monuments, along with understanding of the processes that created them. The efficacy of computer simulation makes it a valuable utility in the analysis of engineering and construction in history; however, a potential point of contention concerns the extent of said efficacy, in that the software requires an optimum balance between data conservation and authenticity. This study demonstrates that although computer simulation via an engineering-gear modeling program can recreate the construction of ancient

structures in accordance to historical records, it is not without its flaws and as such presents important implications regarding simulations of this nature.

2. SOFTWARE SELECTION

Several different programs and/or options can be used to create a simulation of the Colosseum among other monuments. The first option would be to create the model from the ground up, using programs written independently of existing interfaces. However, this approach is disadvantageous, since the matters of rendering objects in the first place must be addressed before any modeling of the Colosseum can take place, and the sheer amount of detail and coding involved would render this strategy impractical.

A less time-consuming method would be to

use an existing graphics program, such as Google SketchUp or Autodesk Inventor, to create the simulation. This approach has the benefit of being far more flexible and faster to implement, because the baseline graphics programming has already been defined.



a)



b)

Figure 1: Two different possible graphics engines were tested for a prototype model of the Colosseum: Google SketchUp (a) and Autodesk Inventor (b). Google SketchUp can integrate the model into an environment, but at the expense of accuracy. Autodesk Inventor, while so far unable to reproduce backgrounds, can produce a more accurate model in the long run, which is why this engine was used for the final model.

Several different graphics programs were considered for use in this study. Originally the intent was to create the model using Google SketchUp and superimpose it over a map of the surrounding landscape. However, the user interface does not support real-time adjustment of

dimensions or automatic feature patterning, and therefore such a task would be fairly arduous at the minimum (Figure 1.a). The most viable option, and the one that is used in the final model, is Autodesk Inventor. Here the model can be shaped more easily and may include a higher degree of complexity for less effort; the disadvantage is that backgrounds are more limited, but the model is intended to provide insight as to how the monument was constructed as well as how it would have appeared, so the surrounding landscape is a minor aspect compared to the model itself (Figure 1.b).

3. RECENT STUDIES

Models of the Colosseum, as with other ancient monuments, do exist and have been created based on historical data. Many of these are reconstructed with varying degrees of accuracy, but what is important is that said accuracy applies mostly to the superstructure because it is the best-known aspect of this monument. Archaeolibri's model of the Colosseum is intended for professional use in archaeological guides, and as such presents an accurate restoration with realistic detail [1]. Another model, produced by Vision Publishing, superimposed digital imaging over a photograph of the modern Colosseum, filling in the gaps left behind by centuries of degradation [4].

One of the major implications of these models is that they show how the monument and surrounding locale appeared upon completion. What is important is that the construction process of the Colosseum, which is of similar importance, is usually overshadowed by the monument's appearance after completion and during its use. For a model that capitalizes on information pertaining to the construction of the monument, the foundation is subject to as much discussion as the superstructure. Starting from this point, the monument can be constructed and then broken down into stages as per the existing data, providing a picture of not only what this building looked like, but also how it was built.

4. MODELING PROCESS

The model used for this project was sculpted, modified, and finalized on Autodesk over at least eight months, partly due to the detailed research that was required to ensure its accuracy and partly due to the complexity of the monument itself. It is based on a combination of the reconstruction in Rome's Museo Colosseo and the 1725 print, *L'Anfiteatro Flavio* [3], modified to fit the dimensions of the existing arena and the exterior walls in the model.

The foundation was of the linear type, meaning the load of the elevation walls was distributed over their planimetric, or two-dimensional, outlining [2]. This means the substructure can be rendered using outlines of the foundation, and can be modeled in a similar manner to the actual building process. An ellipse of the same dimensions as the Colosseum was used as a starting point; the foundation walls were built upon this using extrusion features based on the wall outline (Figure 2).



Figure 2: Foundation construction.

There are two elliptical rings in the superstructure, the inner and outer. To create the level borders and the seats, sweep features, or profiles extruded along a path, were used along with cross-sections of the radial ribs and the building's internal structure, respectively, both obtained from site photos (Figure 3). The path was a projected ellipse derived from the wall extrusion, enabling an elliptical sweep (revolved features have a fixed axis and are used primarily for cylindrical features rather than elliptical ones).

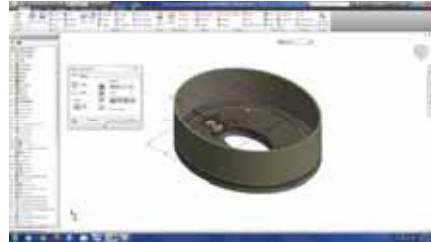


Figure 3: Extrusion features for walls.

The 80 entrances were created using a single Boolean difference extrusion and a procedural array – that is, a pattern of regularly spaced features along a given path. This enabled the creation of three rows of arches, resulting in an 80 by 3 elliptical array of difference extrusions (Figure 4). The windows on the topmost story were created in the same manner later, although only one row was required. The path for the elliptical arrays was later used for the seating sweep feature, as shown in Figure 5.

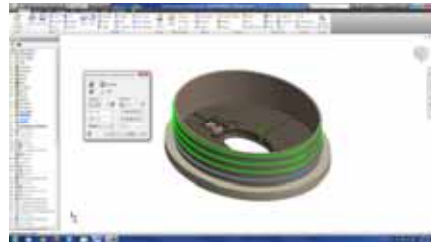


Figure 4: Entrance and archway construction.

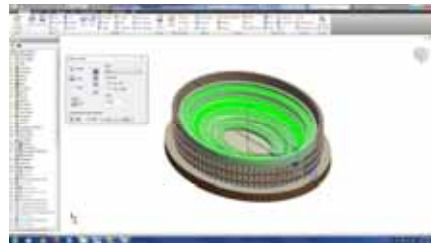


Figure 5: Seat tier construction with Sweep feature.

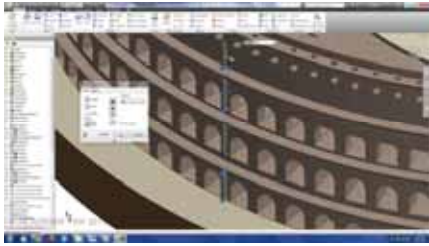


Figure 6: Pilaster construction.

Each of the 80 arches is bordered by a pair of semi-columns; the risk of data lagging limited the detail to a stylized appearance as opposed to a more historically accurate version. A work plane was created at the midpoint between the two arches nearest to the YZ plane. Then four column cross-sections were created on the work plane, and all of them were used in a revolved feature (Figure 6). Finally, another linear array was used to duplicate all of the column features, using the same process for duplicating the archways (Figure 7). On a side note, there were reportedly statues gracing the second- and third-story arches; these were ultimately omitted in this project due to complexity constraints.

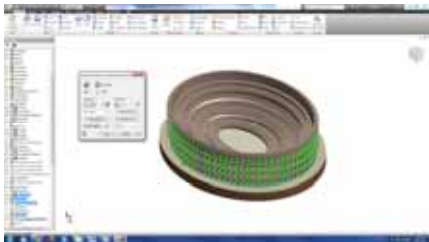


Figure 7: Pilaster patterning.

The passageways leading to and from the seats were recreated via procedural rectangular patterns of extrusion features. The passageways and windows leading to the third-story seat level were replicated through difference extrusions and patterned alternately. The doorways were patterned 40 times, and the smaller windows patterned 80 times, with every other extrusion

not affecting the model due to it extending into the area already removed by the doorway (Figure 8).



Figure 8: Passageway patterning.

Each vomitorium was created using two extrusions: one to create the feature itself and the other to eliminate the stairs inside (Figure 9). These were elliptically patterned per the quarter-section view from Carlo Fontana's 1725 print, *L'Anfiteatro Flavio* [3] (Figure 10). The stairways were created using a different approach because a rectangular pattern would not have lined them up evenly between the vomitoria. Five stairways were created independently with five separate work planes based on axis points between each of the vomitoria. Then these stairways were mirrored on both the XZ and YZ planes to create 20 different stairways.



Figure 9: Vomitoria construction.

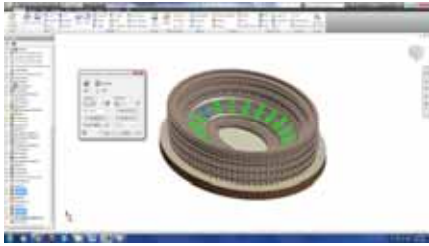


Figure 10: Vomitoria patterning.

The final component of the superstructure was the velarium, which has left the least physical evidence. A common theory was that it was hung on a network of ropes tied to the 240 masts that would have fit into the corbels of the structure; however, other evidence suggests horizontal booms supporting a shorter awning (Figure 11). These two possibilities were reconstructed as an experiment by a team of experts under the surveillance of NOVA, the science television series. It was deduced that the mast system would have been more likely (Figure 12); though shorter, it would have been easier to retract, and was surprisingly effective as a sun shade [5].



Figure 11: A close-up of the fresco, *Brawl at the Pompeii Amphitheater*. Note the velarium structure outlined in red, over the top of the stadium. (Robert Etienne: *Pompeji, die eingeäscherte Stadt*, Ravensburg 1991. Museo Archeologico Nazionale, Naples.)



Figure 12: Two possible models for the *velarium*. The rope-supports are more vulnerable to high winds, while the masts are shorter but can be retracted in hostile weather (WGBH 2006).

Due to data constraints, a simplified version of the velarium was created for the project. The 240 corbels and the corresponding holes on the topmost exterior ribs were replicated via the same rectangular pattern approach used for the archways. The masts and booms were created with two revolved features (i.e. revolving a profile around a central axis) that were copied likewise. The velarium is an extruded feature, but with a nominal thickness (about .1 m or so) to pass off as a membrane, as surface features are typically translucent/transparent (Figures 13, 14).



Figure 13: Colosseum without *velarium*.

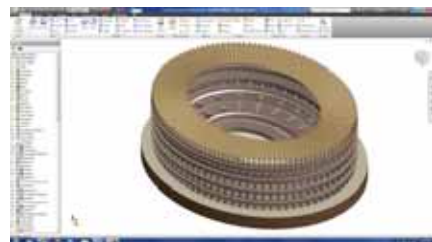


Figure 14: Colosseum with *velarium*.

The final phases of creating the simulation, the level stages and section view, were created from the finished model. The two main processes involved are respectively to subdivide the model into individual levels and to cut a section from the model before replicating the interior. The superstructure was divided via extrusion features that removed each of the stages, starting from the attic down to the ground floor (Figure 15). This is where the disadvantage with the method described became clear; the model was created in one solid piece, meaning that replicating the individual sections would require a more powerful engine to accommodate the increase in complexity.

For the section view in Figure 46.a, a segment each of three revolved difference features was positioned so that the resultant slice could provide a cross-section of the model's interior without interfering with the seating. With the features stacked over each other, the end result looks similar to Figure 46.b at first glance. However, a closer examination reveals that in comparison to the scale model, the interior of the digital model is not accurate, primarily because the stairways were not modeled during the initial sweep feature that created the interior and seating (Figure 16).

To remedy this for the section view, the stairways were created manually for each side of the split. In the figure, the right side has two stairwells, one on top of the other. These were created using a single sketch feature, which was marked with the Shared attribute meaning that the sketch could be used to create multiple features. The stairways are drawn in by hand, with excess line-work trimmed and open loops closed as necessary. Another set of extrusion features is used to create the stairways (Figure 17).

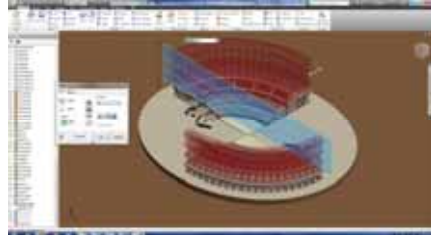


Figure 15: Reverse-engineering the model construction stages.



Figure 16: The inaccurate sectional view prior to modification.



Figure 17: Extrusion of the dual stairways.

The side of the section view that had two stairways in opposite directions connected by a landing was more difficult to construct because it required some modification of the cross-section to accommodate them. A series of extrusion features was used to clear out the protruding portions of the interior. Notably, the connected stairwells were created from difference extrusions because it was concluded that excess system lag could be reduced by creating them from the existing wall (Figure 18).



Figure 18: Removal of the filled interior space, to make way for the connected steps.

The completed model is shown in Figure 19. Notice that to complete the image, a ground object was added and preprogrammed shading was applied for realism.



Figure 19: The completed Colosseum model, with perspective, preprogrammed shading and a ground platform.

5. ANALYSIS

The stages of the Colosseum erection as recreated by the model are shown in Figures 20 through 24. Two of the four sections have been removed, so the cross-section of the Colosseum during each stage is also visible. A number of views of the Colosseum model can be used to judge the integrity and reliability of the model in comparison to the actual monument. At first glance, the exterior is similar to other restorations of the monument in terms of the outer décor, and the interior shows the vomitoria and stairwells in positions akin to the historical print used as a reference. However, despite resemblances to the actual structure, there are subtle differences which could potentially compromise the accuracy of the model.



(a)



(b)

Figure 20: The foundation prior to the first level of construction (a) without the floor surface/arena and (b) with the floor surface/arena.

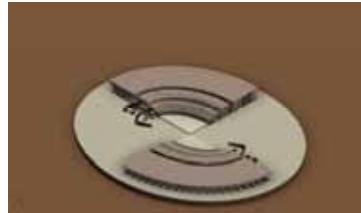


Figure 21: Cross-sectional view of the first level of construction.



Figure 22: Cross-sectional view of the second level of construction.

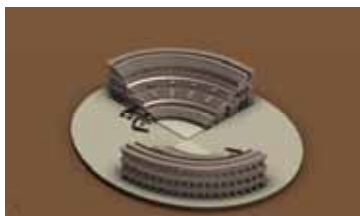


Figure 23: Cross-sectional view of the third level of construction.



Figure 24: Cross-sectional view of the fourth level of construction.

For example, the floor of the arches is not on the same level as the radial ribs, but above them because of the different distances for the rectangular patterns (Figure 25). Inside, the vomitoria are dead-ends because the stairways overshadow the passageways into the interior. Additionally, the second floor has raised areas between the archways, and the first floor is level to the ground plane.

The contrast between the interior of the digital model and the interior of the scale model shown in Figure 29 shows the contrast between the two models more clearly in the second animation. The initial zoom-in is directed towards the foundation of the model, which is obscured by the floor. This can be justified by Inventor recognizing additions in a part as union features, which is particularly important when the fact that this model is essentially one large, complex part is taken into consideration. And anyway, this is the least of the setbacks demonstrated in this animation.

The stairways in particular are not as well designed as other aspects of the model. The interior of the model is created entirely using

sweep features to reduce complexity, but this is at the cost of accuracy. The interior would have had a complex network of stairs and passageways, but the paired stairways of the model section view run into the tops of the columns in front, and the walled stairway does not have entry or exit archways because it is too narrow to accommodate these.

6. DISCUSSION

As the modeling process demonstrates, this simulation can allow for relatively complex details such as the seating, vomitoria, and outer décor with simple extrusions, sweeps, and patterns. The model was constructed without having to divide or mirror a fraction of the component because the rectangular arrays can factor in an elliptical path, with the orientation of the components adjusted to the path. Likewise, sweep features can be used to render elliptical features such as the seat tiers and internal structure with a relatively simple series of steps. In theory this degree of complexity can be achieved with any similarly advanced graphics software provided that it has the functional capability of handling the large number of features required. The most significant problem with the modeling simulation would be handling the complexity, due to the sheer amount of data the system must process. At higher levels of complexity, if too many steps are taken at once, the program may stall or, in the worst case, crash altogether. This means the computer hardware requires further development to catch up with the software requirements and capabilities in order to make this simulation viable. The overall result is a trade-off between rendering capabilities and accuracy, which explains the relative simplicity of some of the more minor elements of the model, such as the outer pilaster caps.

Comparing some of the specific features of the model and the corresponding sites in the Colosseum, several further setbacks are revealed. Based on a historical source, the model does simulate the exterior adequately, but the

same cannot be said with respect to the interior. This is especially apparent in the comparison shots shown in Figures 25 through 29; the model features are solid, monolithic structures without subdivision, and overlook key structural components such as keystones in archways, capitals in columns, and so on. So the model is not truly indicative of the appearance of the monument prior to substantial damage and/or renovations. That said, however, the methods of such a simulation can vary enormously. Given a more powerful engine, better software capabilities, and a more comprehensive data pool, it may be possible to recreate the Colosseum interior, but this would come at the expense of both accessibility and comparative ease of use.

The disparities shown in the previous pages suggest that depicting the construction process may be even more taxing than depicting the model itself. The processes for building the Colosseum demonstrate different methodologies with regards to the way the monument was erected, with the digital model being constructed starting with the full walls rather than in levels like the original monument. This is not to say that the ancient construction methods cannot be rendered digitally; it is simply less time-consuming but also less inaccurate to complete the structure before working the steps in reverse, based on the limited historical data that can be gleaned with regards to these methods.



(a)



(b)

Figure 25: A column abutment from the Colosseum interior (a) is compared with a corresponding structure from the model (b), which shows that the model column is relatively unrealistic, partly due to the lack of photorealistic texture.



(a)



(b)

Figure 26: A quick comparison between one of the vomitoria in the Colosseum (a) shows that it does not have the attic overhead, as in the corresponding structure in the model (b).



(a)



(b)

Figure 27: The arena of the Colosseum has degraded over time (a), and the same can be argued for the concrete that comprised most of the seating. However, the model (b) used a historical diagram as a basis, and may provide some ideas of the monument's appearance in its still pristine state.

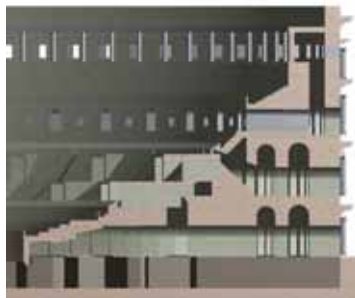


(a)

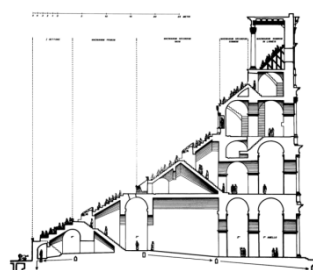


(b)

Figure 28: A sample archway leading up to the seating (a) is further corroboration of the model's interior underdevelopment, as it clearly shows an opening towards the seating which, in the model, is covered by the stairway just behind it (b).



(a)



(b)



(c)

Figure 29: This series of cross-sections shows how the Colosseum model can reflect inaccuracies that may result from subjective viewpoints regarding the reconstruction. Miscalculations in the model mean that the seating in (a) has a lower slope than in the official cross section from the Museo Colosseo (b) and in the section scale model from the same museum (c).

7. CONCLUSIONS

Digital simulations of ancient monuments such as the Colosseum may provide a lucrative outlet for such complex structures that can be reconstructed with enough reference material and a bit of speculation. The Colosseum in particular has significance with regards to construction and architecture, with a number of sources providing historically accurate dimensions which can be used for the pictured simulation. And the Romans used a blend of several different styles along with innovations of their own to develop their own fashion of construction and engineering which remains influential. As such, it is expected that the Colosseum can be created in graphical form with enough detail to remain recognizable. Constructing the simulated model, like the actual monument, starts from the ground-up with a foundation built in a similar fashion to source records, while the superstructure uses processes that simplify construction but rely on dimensions with little regard for historical methods.

Where this idea falls flat, in the general sense, is in the simulation of the construction process. The procedures used to create the virtual model in the first place are not without their setbacks, the most notable of which is the increase in render time relative to complexity. More importantly, while this approach could work in theory, the very process of reverse-engineering the procedures could provide an entirely new set of problems, such as subdivision of the structure, recreation of individual features, and possible scaling of equipment and other utilities.

Despite the potential shortcomings of visual simulations of ancient construction, future research and development is nonetheless encouraged in this area of study. Advancements in both computer graphics and historical research may be beneficial in the virtual reconstruction of monuments such as the Colosseum. In due course, such reconstructions can provide an increasingly detailed look into the era when the respective structures were built, thereby en-

hancing public and academic knowledge of both these ancient monuments and their respective areas and time periods.

ACKNOWLEDGEMENTS

The author dedicates this paper to his parents, whose love, encouragement, and support have allowed the research required to be possible.

REFERENCES

- [1] Coletta, Giuliana. *Rome Reconstructed*. Archaeolibri, 2010.
- [2] Malacrino, Carmelo G. *Constructing the Ancient World: Architectural Techniques of the Greeks and Romans*, pages 120-122 J. Paul Getty Museum, 2010.
- [3] Quennell, Peter. *Architecture of the Western World* pages 38-39. New York: Newsweek, 1971.
- [4] Staccioli, R. A. *Rome, Past and Present*. Vision Publishing, 2001.
- [5] WGBH (Television station: Boston, Mass.), British Broadcasting Corporation., & WGBH Video (Firm). *Secrets of Lost Empires: Colosseum*. Boston, MA: WGBH Boston Video, 2006.

ABOUT THE AUTHORS

1. Adrian H. Tan is a graduate student at the Ohio State University. Adrian has a B.S. in Computer Science and Engineering and an M.S. in Civil Engineering from the Ohio State University. Adrian is currently working towards a Ph.D. in civil engineering and construction with a focus on computer graphics and virtual simulation in the engineering industry.
2. Dr. Frank M. Croft Jr., P.E., is an Associate Professor in the Department of Civil and Environmental Engineering and Geodetic Science at The Ohio State University, where he has been since 1984. Dr. Croft has a B.S. in Aerospace Engineering from the Indiana Institute of Technology, an M.S. in Engineering from the West Virginia Col-

lege of Graduate Studies, and a Ph.D. in engineering at Clemson University in South Carolina. Dr. Croft has been very active with the Engineering Design Graphics Division (EDGD) of the American Society for Engineering Education (ASEE) since 1974.

3. Dr. Fabian H. Tan is a Professor in the Department of Civil and Environmental Engineering and Geodetic Science at The Ohio State University, where he has been since 1982. Dr. Tan has an M.S. in structural engineering, an M.E. in construction engineering and management, and a Ph.D. in construction engineering and management from the University of California in Berkley. Dr. Tan has served as a forensic consultant with numerous law firms in Ohio as well as the United States Air Force Weapons Laboratory in Kirtland, New Mexico. Additionally, Dr. Tan has also served as a project manager and engineer on multiple international construction projects such as Leighton Australia-Indonesia Construction Inc. and Mahkota-Ekman Sweden Inc. Dr. Tan is currently teaching History of Ancient Engineering at The Ohio State University.

SOME CASES USING DESCRIPTIVE GEOMETRY IN VARIOUS COURSES

Shinobu NAGASHIMA
Rikkyo University, Japan

ABSTRACT: This paper discusses the use of descriptive geometry in various courses. It is important to imagine and understand figures and shapes in three dimensional space. But it is not easy to learn about three-dimensional geometry. In Japan, descriptive geometry is taught for mechanics, architecture, civil engineering and so on at many universities for these 100 years. These days CAD took the place of design and drawing, descriptive geometry is going out year by year. I think it is important to continue teaching descriptive geometry, and I tried to teach descriptive geometry in various courses, but educational estimation has not yet done.

Keywords: graphics education, descriptive geometry, mathematics education, computer graphics.

1. DESCRIPTIVE GEOMETRY

Descriptive geometry was discovered in 18th century by Gaspard Monge[3]. Descriptive geometry was taught as a basis of mechanical design, but it was replaced with CAD these days. It is important to imagine and understand figures and shapes in three dimensional space. But it is not easy to learn about three-dimensional geometry. The role of descriptive geometry is to help for understanding and solving 3D problems. I think it is important to continue teaching descriptive geometry[4-6], and I tried to teach descriptive geometry in various courses.

2. MATHEMATICS

First, I show you a topic of conic sections in mathematics. Figure 1 shows intersection of a cone and a plane, these figures show the intersection curve becomes ellipse. The figure is from the "Dandelin spheres" in Wikipedia[2], we often see this typical figure in usual explanation.

Two spheres G_1 and G_2 are touching inside the cone and touching the plane at F_1 and F_2 . These two points are called foci. The point

P is an arbitrary point on the ellipse. The length $PF_1 = PP_1$, and $PF_2 = PP_2$, so $PF_1 + PF_2$ is equal to $PP_1 + PP_2$. This length is constant. It is not easy to know a sphere G_2 is really touching the cone and the plane, or the G_2 exists uniquely.

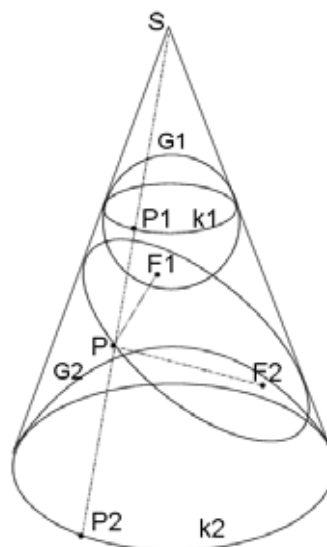


Figure 1: Cutting of a Cone.

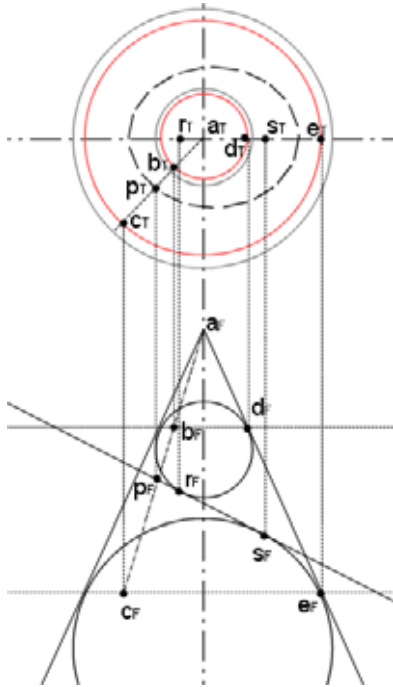


Figure 2: Cutting of a Cone (ellipse).

Figure 2 was used in a course of mathematics, analytic geometry at Rikkyo University. Front and top views are drawn and we see the plane as a line (edge view). The small sphere is touching inside the cone at B, D the big one is touching at C, E and the two spheres are touching cutting plane at points R, S also. Then $BC = DE$, $BP = PR$, $PC = PS$, so we know $BP + PC = BC$, $BC = PR + PS = DE$.

The distance $PR + PS$ is equal to DE (constant). We know two spheres exist uniquely, touching the cone and the plane.

It is hard to calculate these distances by coordinates and it is hard to understand this fact by reading such explanation. But it is not so hard to understand it by thinking with this descriptive geometry figure.

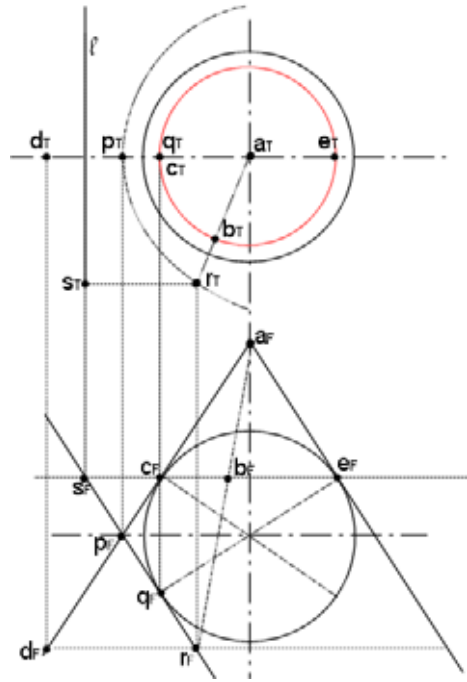


Figure 3: Cutting of a Cone (parabola).

Figure 3 is a special case of conic section. Namely the cutting plane is parallel to AE . The intersecting line between cone and a plane is a parabola. In this case, the line l (passing point S) is called directrix. Only one sphere is touching inside the cone and touching the plane at point Q , which is called a focus. Point R is an arbitrary point on the parabola.

We know $RQ = RB$, because two lines are touching the same sphere. And we know $RB = CD = RS$, so $RB = RS$. This is an important property of parabola. The length $R_F - S_F$ in front view is a real length of RS in space, we know easily by using descriptive geometry.

3. GRAPHIC SCIENCE

3.1 Astronomical topic

This topic is used in graphic science course at University of Tokyo. Table 1 shows sunrise and sunset time in Japanese three points at winter and summer solstice[1]. These three points are Nemuro, Tokyo and Naha which are located top, middle and down of Japan Islands (Figure 4).

Figure 5 shows rotation of the earth at winter solstice. The right side of the earth is day, and the left side is night.

Table 1: Sunrise and Sunset time in Japan.

	winter solstice		summer solstice	
	21 Dec 2013		21 Jun 2014	
	sunrise	sunset	sunrise	sunset
Nemuro	6:46	15:45	3:37	19:02
Tokyo	6:47	16:31	4:25	19:00
Naha	7:12	17:42	5:37	19:25
max difference	27 min.	117 min.	120 min.	25 min.

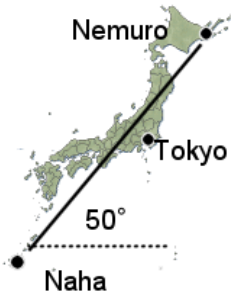


Figure 4: Japan’s Shape.

Front view in Figure 5 shows sunrise moment. The rotation axis inclines 23.5 degrees, and Japan’s shape is almost a bar inclined 50 degrees. So Japan’s bar inclines $23.5+50=83.5$ degrees against horizontal line, it is almost vertical. Sunrise times of three points are almost same.

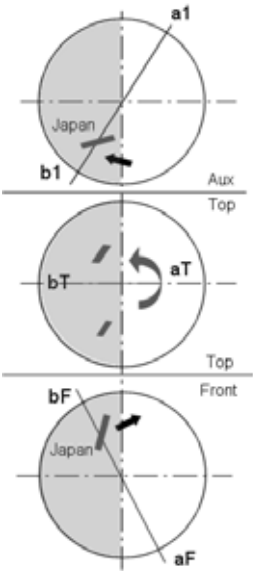


Figure 5: Rotation of the Earth (at Winter Solstice).

Top view shows rotation of Japan, about 180 degrees rotated means sunset position. Back view shows at sunset time, Japan’s bar is nearly horizontal, so sunset times are rather different.

In Summer solstice, day and night are changed. The sun and the earth positions are reverse. Maximum difference of three points’ sunrise time is long, sunset time is short as shown in Table 1.

3.2 Topics of Polyhedron

In graphic Science Course, we often draw regular polyhedron. By thinking construction of a body from triangles or pentagons, we know symmetry and regularity naturally. We can see a cube in a Dodecahedron and three axes in an Icosahedron as shown in Figure 6.

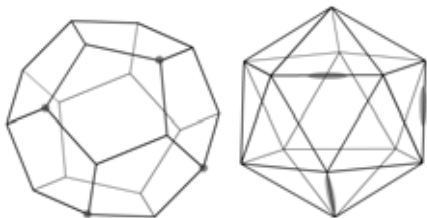


Figure 6: Dodecahedron and Icosahedron.

Figure 7 shows orthographic views of a cube. It is easy to draw these views so all students may draw these views. Next, I gave students an instruction to transform the cube.

(1) Lift up the center of a face and make a square pyramid.

(2) Make six square pyramids on every face of the cube.

(3) The height of a square pyramid is determined to create a rhombus form next two triangles (Figure 8). This polyhedron has twelve rhombuses, and this is called “Rhombic Dodecahedron”.

Many students can draw this polyhedron, but they cannot understand that this polyhedron is a part of right hexagonal prism. I made students to make a paper model of this polyhedron (Figure 9), and I showed this polyhedron has a space filling property (Figure 10).

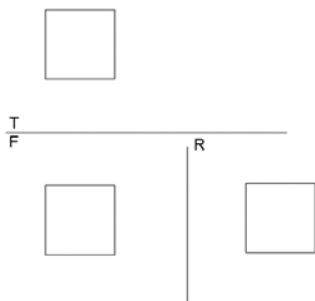


Figure 7: Cube.

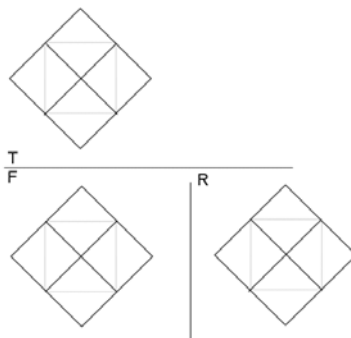


Figure 8: Rhombic Dodecahedron.



Figure 9: Paper Models of Rhombic Dodecahedron.



Figure 10: Space Filling Property of Rhombic Dodecahedra.

4. COMPUTER GRAPHICS

Figure 11 shows output of cloth model using POV-Ray software in computer graphics course at Bunka Gakuen University. This cloth consists of two types fibers. Fibers consist of thousands of spheres, and we have to determine coordinate of the center of all spheres. I designed the shape not to intersect two fibers by using descriptive geometry (Figure 12).

It is not easy to know two fibers are intersected or not. But if fibers are intersected, it has no problem, POV-Ray can display by ray tracing method.

To solve curved line or curved surface problems are difficult in descriptive geometry. In this case, it is important to have knowledge of primitive objects, such as circle, ellipse, sine curve and so on.

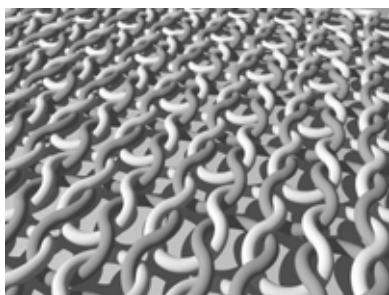


Figure 11: Fiber Model by POV-Ray.

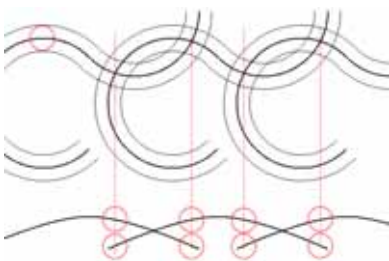


Figure 12: Design of Fiber Model.

5. CONCLUSIONS

Some cases using descriptive geometry were shown in this paper. I showed some examples of descriptive geometry, in mathematics, graphic science, and computer graphics courses at three universities. Descriptive geometry often has a special power which gives us understanding way for various problems. It is not easy to gather and teach these descriptive geometry's methods in one course. Though I did not estimate the merit of descriptive geometry, I hope descriptive geometry is used when it is needed.

ACKNOWLEDGMENTS

I would like to thank Professor Kenjiro Suzuki, National Institution for Academic Degrees and University Evaluation.

REFERENCES

- [1] <http://eco.mtk.nao.ac.jp/koyomi/dni/>
- [2] http://en.wikipedia.org/wiki/Dandelion_spheres
- [3] V. J. Katz. A history of mathematics: An introduction (2nd edition). Addison Wesley, 1998.
- [4] S. Nagashima. et al. Development of software for teaching descriptive geometry using authoring system, *Proceedings 11th ICGG*, pp.516-519, 2004.
- [5] S. Nagashima. Development of the data-base system for descriptive geometry and its application, *Proceedings 12th ICGG*, 2008.
- [6] H. Stachel. The status of today's descriptive geometry related education (CAD/CG/DG). *Proceedings of JSGS Annual conference*, pp.15-20, 2007.

ABOUT THE AUTHORS

Shinobu Nagashima is a professor, at Department of Mathematics, college of Science, Rikkyo University.

SPACE DESCRIPTION PROBLEMS IN THE SECONDARY SCHOOL (THEIR POSSIBLE CAUSES AND PROBLEM SOLVING)

Rita KISS-GYÖRGY

Belvárosi I. István Secondary School, Székesfehérvár, Hungary

ABSTRACT: I have been teaching Art and Design and Technical Description in some classes. I have a degree in Art, an MSc in Mathematics and an MSc in Descriptive Geometry (in Debrecen). I have experienced that space description, the position of objects in space and the 2 dimension-mapping and drawing them makes it more and more difficult. This problem occurs in the Technical Description classes and last but not least in the Math lessons as well, mainly at the topic of space-geometry.

In this thesis I want to give a short review of the Technical Description and Art topics before the appearance of the National Curriculum - what the students had already studied or they should learn.

I am going to present some typical problems of spatial and art description and editing mistakes.

I am going to analyze - with tests and questionnaires - how strengths and weaknesses of space sight and description are influenced in everyday life. I am trying to explore the roots of the problem based on the results and to detect the causes of typical problems.

I am going to present the topics of the new National Curriculum introduced in 2013 – the number of Technical Description and Art lessons have dramatically decreased.

Finally, I would like to find and create some effective methods for the secondary school students to improve their space description and space sight abilities during the lessons based on the results of my survey.

Keywords: space description, typical problems, Art, Technical Description, National Curriculum.

1. INSTALLATION

The problems

Spatial vision and an adequate space perception are essential to everyday life. Despite of these, there are very few possibilities in the school curriculum. For example, fewer and fewer Art lessons, and it is already moving towards digital applications, the geometry involved in mathematics education is neglected, and in the professional subjects of descriptive geometry, technical presentation does not get enough number of lessons.

Although the map view should be developed continuously. Unfortunately, eliminating the space geometric tasks begins between the ages of 10 and 14 in the elementary school, – for

example: the neglectation of spatial representation in sketching and spatial modelling does not appear in none of the subjects (and due to the lack of time in secondary education a three-dimensional construction and model-making are missing).

I have experienced that space description, the position of objects in space and the 2 dimension-mapping and drawing them makes it more and more difficult. This problem occurs in the Technical Description classes and last but not least in the Math lessons as well, mainly at the topic of space-geometry. While the role of mathematics as the basis of all scientific knowledge, and geometry as the heart of mathematics (Mlodinow, 2001).

The representation of space is becoming less

advanced among students. Their drawings are less expressive, more and more do not display spatiality, they represent the spectacle in less and less unified way from their viewpoint.

I would like to find out why this occurs, and why students between15-18 draw in a less expressive way.

The advancement of technology such as computer graphics and multimedia have greatly expanded the effect and power of visualization in every field. Because of the spread and popularity of 3-D movies, televisions and computer games we would expect from our students to have a more developed spaceview than the older generation.

2. PARTICIPATING CLASSES IN THE SURVEY:

There are ten classes I have chosen for my analysis (Table 1 and Figure 1).

Five classes have already studied or still study Technical Description, and five classes have not studied it in secondary school.

Among the surveyed students there are 52 girls and 41 boys who have studied, and 72 girls and 23 boys who have not studied any Technical Description.

Eight classes have already studied or still study and two classes have not studied Art in secondary school yet.

Table 1: Surveyed students

Year 2013-14	Studied Technical Drawing					Not studied Technical Drawing				
	No Art	Learned / learn Art				Learned / learn Art				
Classes:	9.C.	9.D.	11.C.	12.D.	13.C.	10.A.	10.B.	10.C.	10.D.	11.B.
Total number	39	30	27	17	23	27	27	25	22	32
female	19	19	20	8	12	18	23	19	10	26
male	19	11	7	9	11	9	4	6	6	6
Surveyed (people)	25	21	16	13	18	20	17	21	13	24
female	8	13	14	6	11	14	16	15	7	20
male	17	8	2	7	7	6	1	6	6	4

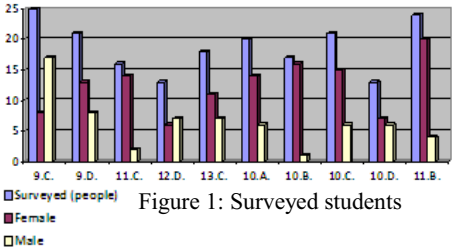


Figure 1: Surveyed students

3. OUTLINING SYLLABUSES

If we take a look at the Technical Description and Art topics before and after the appearance of the new National Curriculum - that the students had already studied or they should learn – and compare the two themes, and their number of hours, we could see, the number of Technical Description and Art lessons have dramatically decreased!

They will have 17 hours instead of 72 hours (or in vocational training 36 hours).

There are many problems in the Technical Description lessons:

The first lesson of the school year has not yet been installed and there are no editing tools!

There is no textbook, so I have to dictate everything, and I have to pay attention to everyone by editing, and it takes a very long time.

Due to the reduced number of hours, I wrote a note which contains the theoretical and practical elements. This can be divided into lessons, and can be downloaded and printed from my website. But, unfortunately, students do not watch or print the material (many of them do not have a printer at home, and we cannot make it compulsory). There is no way to copy 3-4 pages for everyone in the school.

The number of Art lessons decreased too: 36 hours instead of 72 hours (and in some schools, there will be no visual culture hours at all, was selected as one of four subjects – Drama, Media studies, Singing and music and Visual culture.) Topics, according to the New Curriculum are absolutely different. In the Old Curriculum the theoretical and practical parts follow each other in proportion, as well as some topics appear cyclically. In the New Curriculum the topics are in blocks.

Problems in Art lessons:

In the first lesson of the school year students do not even have equipment!

They did not use to have and they still do not have a textbook.

Technical Description appears in the new

subjects (acquiring of the basics), then they have to use them in the design. I think, because of the slenderness of time, this is not realistic at all.

The new curriculum would require digital equipment: computer, camera, camcorder - but such amount (minimum 24 persons / class; and there are 4 classes) are not available. So the requirement is impracticable because of the class size. For example, the old themes, mainly of art history section I often need a laptop, projector and an extension cord, but there are some institutions where you do not even this type of equipment is provided.

The new curriculum requires a background knowledge of the students in art history and digital competences students acquired in elementary school. The students who learned according to the old syllabus have very little basic knowledge of art history.

A new framework assumes that students can draw figures, objects and space in expressive way. The emphasis is not on the freehand drawing, but the design, where they would apply the acquired knowledge which comes from the elementary school. (I have my doubts about it - which is supported by my experience.)

I wonder, without the appropriate manual skills and abilities, how this can be carried out. (Only time will tell whether we will be successful in the years to come.)

4. TYPICAL PROBLEMS AND MISTAKES IN EDITING AND DRAWING TASKS

4.1 Typical mistakes in Technical Description tasks

Generally speaking, the thread of editing, editing accuracy, the different views, the relationships among arranged images, and the imagination the 3-dimensional object from its images are the main problems.

Why is it necessary to edit exactly? Why it is not indifferent if two lines parallel or not (for example the two bases of a trapezoid)?

The projection and the axonometric representation (pictorial representation) have already created problems. How are the axis situated? Which size should assess in which axis? How will pyramid be straight? How to measure the height of the pyramids? (Figure 2 and 3.)

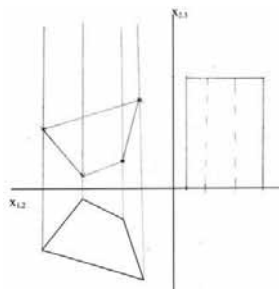


Figure 2: Column in Monge-projection
(9.D. girl)

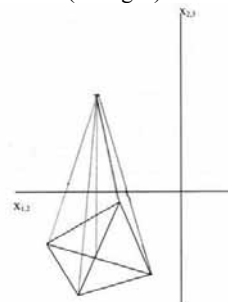


Figure 3: Pyramid in Monge-projection
(9.D. boy)

4.2 Some typical problems about spatial representation in Art classes

Lots of students take a photo about a view (about a still life) and make a picture - a copy - from it. To translate the three dimensional view to a paper is bigger problem than copy reproduction, so the space mapping to the plane is one of the main problems.

Students are very unimaginative. The other main problem is that they can not draw their ideas in an expressive way. How the pattern follows the shape of a vase (in a solid of revo-

lution)? (Figure 4 and 5). Some of the students are unable to interpret which part is not visible.



Figure 4-5: Greek vases (10.C. girls).

Some of them do not understand, why building cannot start in the horizon line, why should "hang well below" the base plane.



Figure 6: Room from memory (11.B. girl)



Figure 7: Room from memory (10.C. girl)

Only a few students try to apply perspective on their work. Often mixed an axonometric and planar elements (Figure 6 and 7), appears the mixed view point representation, and sometimes also a perspective view or axonometric view of the objects. Many of them find it difficult to "narrow" the room into spectacular pieces.

5. INTRODUCTION OF THE QUESTIONNAIRE AND THE TEST

I began to consider and analyze in school year 2013/2014 - with tests and questionnaires - how strengths and weaknesses of space sight and description are influenced in everyday life. I am trying to explore the roots of the problem based on the results and to detect the causes of typical problems.

5.1 About the questionnaire

According to the results of an experiment conducted in 1984, the development of space perception is significantly influenced by three factors: the person's sex; grade from Mathematics and grade from Art (Drawing) [1].

As a fourth factor I am trying to compare the results with the Physical Education grades, because good spatial coordination and perception are essential for good sport skills.

As a fifth moment I pay attention games to played in students' childhood, what kind of plays they liked, or disliked. I suppose that the traditional building blocks, construction toys and LEGO significantly develop children's spatial vision [2].

Another assumption is that the three-dimensional computer games are developing space vision. They are obviously useful, but I think, shooting, clicking and pursuing activities in the 3-dimensional virtual space, do not substitute actual construction in real space.

5.2 About the test [3]

Primarily, the test contains tasks for analyzing students' spatial vision.

Some tasks require basic mathematical and geometrical knowledge - for example, the

search for different rectangles, knowledge of the triangle types. Other tasks require some imagination - for example imagine the spatial objects, and find the opposite face of the cube.

In the test students can obtain 44 points with the exception of task 1.b.) - because there are up to 32 points which can be collected. In class 9.C. there were two girls (Figure 8) who scored a lot of points in that task. (They sat next to each other.)

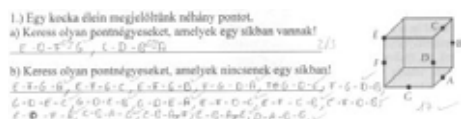


Figure 8

The average points of the test, and the most and the least points in classes (Table 2; Figure 9).

Table 2: Average points of the test

School year 2013-14	Had Technical Description					Had not Technical Description				
	No Art	Had	have Art			Had	have Art			
Classes:	9.C.	9.D.	11.C.	12.D.	13.C.	10.A.	10.B.	10.C.	10.D.	11.B.
Number of participants:	29 p.	21 p.	18 p.	17 p.	18 p.	20 p.	18 p.	21 p.	17 p.	24 p.
Number of participants:	93 people (students)					64 people (students)				
Average scores on the test	19	18.9	16.25	18.89	17.44	14.68	13.82	18.714	17.907	17.218
The most scores	36	31	34	32	39	27	21	29	32	29
The least scores	6	6	6	10	9	7	8	8	8	8

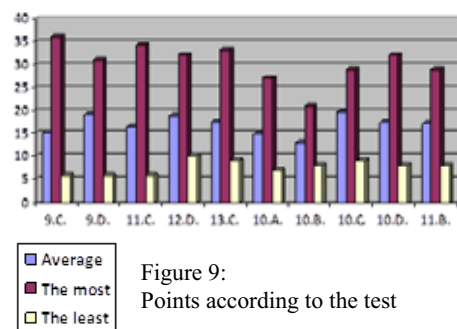


Figure 9:
Points according to the test

6. RESULTS OF THE SURVEY

(Analysis of the questionnaires, and comparing the marks of Mathematics, Physical Education

and Drawing for the test results is still ongoing.)

Task 1.

36.5% of students who have learned Technical Description, given two planar rectangles, 23.4% of students who have not learned Technical Description, given two planar rectangle, but there were two people (1.06%), who found three two-dimensional rectangle.

In class 10.A. no one had two or more solutions.

Among girls who have studied Technical Description, performed better 15% the average and the boys who have spent Technical Description also performed better, although the difference is not so big.

The difference among girls who wrote two good solution (in%) is less. The girls who studied Technical Description were 9% more two correct answers.

The boys who have studied Technical Description answered two good solutions more than twice as many - in% - (39.02%) than boys who have not studied Technical Description (18.18%).

However, only one boy did not learn Technical Description gave the third two-dimensional rectangle, and he only listed three rectangles.

Task 2.

It is shocking that the students are not able to name and identify the types of triangles according to sides and angles. And the maximum of 10 points for most of them (69 people) have achieved only two points, so the 36.89% of the students who completed the test (a bit more than 1/3 of them)!

Only five people called (one or two) triangles according to the right sides and angles of all surveyed students (187 people). They all learned Technical Description.

The students who had Technical Description (93 people) have learned it. (Of course, all students should know, as they all learned or learn-

ing Mathematics.)

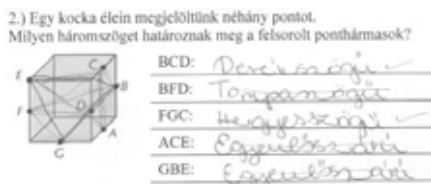


Figure 10: Type of triangles
(12.D. girl)

35.3% of surveyed students – 32.2% who had Technical Description; 38.3% who had no Technical Description; 44.3% of females and 12.7% of males - wrote that the BFD triangle is an obtuse triangle (Figure 10 and 11)!

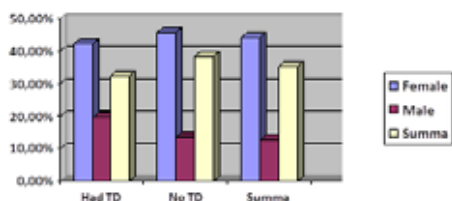


Figure 11

Task 3.

Only 8 students (out of 187 people) reached 4 points out of the maximum 7 points (that is 4.24% of the students who wrote the test), 2 of the 8 students (1 boy and 1 girl) have already studied Technical Description, and 6 students (4 boys and 2 girls) have not studied Technical Description before.

They wrote 1-1 different answers to each quadrilateral, and there was nobody who had known, that a rectangle and a parallelogram is a trapezoid and a parallelogram is a trapezoid too.

In class 13.C. there was 1 boy who knew/wrote that a parallelogram is a trapezoid too – he had studied Technical Description before.

In class 11.B. there was 1 girl who

knew/wrote that a rectangle is a trapezoid too – he had not studied Technical Description before.

The difference is not so significant among students who studied or did not study Technical Description, with the exception of the deltoid, in which students who have studied Technical Description gave approximately 5% more correct answers.

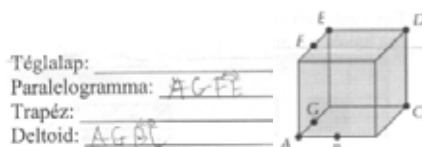


Figure 12: Rectangles
(9.C. boy)

Most of the good answers have been given to the parallelogram by 87 people out of 187 members; that is 46.52%.

The second most of good answers have been given to the deltoid by 78 people out of 187 members; that is 41.71%.

The third most of good answers have been given to the rectangle by 51 people out of 187 members; that is 27.27%.

The least of the good answers have been given to the trapezoid by 30 people out of 187 members; that is 16.04% - despite the fact that there were three correct answers.

Almost 1/3 of the surveyed students by 59 people out of 187 members - that is 31.55% - did not write, or did not give a correct answer!

Task 4. and 5.

Generally, more students drew the projections correctly about the body, (although nobody drew the arranged fundamental dimensions – this is influenced by the previous task, as there were currently no arranged fundamental dimensions of each view.)

The students who have studied Technical Description could draw the image of the body in higher proportion (46.23%), compared with the students who have not studied Technical

Description (31.91%). The difference is about 15%!

However those who have not studied Technical Description performed better at drawing a picture of the body of its projections, (61.7%), than those who have studied Technical Description (53.76%) before. The difference is about 8%.



Figure 13: Task 4 and 5 from the test (10.A. girl)

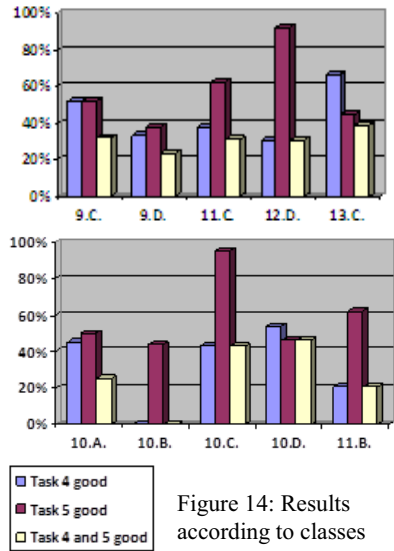


Figure 14: Results according to classes

The students who gave good solutions for both Task 4 and Task 5 have not studied Technical Description (30.1%), while those who have studied Technical Description, 26.59% of

them had good solutions. (The difference is not so big, about 4%.)

Results according to gender: around 9% of boys performed better in Task 4. and 5. than girls.

Task 6.

Most of the girls who have studied Technical Description before, usually gave two good responses (42.3%).

Most of the girls who have not studied Technical Description before, usually gave one good response (36.11%).

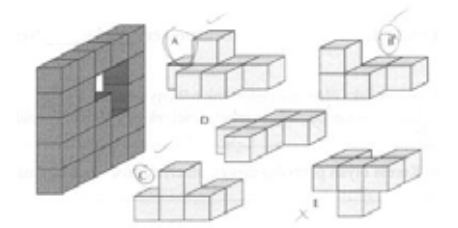


Figure 15: Three good responses out of four (9.C. boy)

Most of the boys who have studied Technical Description before, usually gave three good responses (39.02%) (Figure 15).

The boys, who have not studied Technical Description before, gave one, three and four good responses in equal proportions (27.27%).

Those who marked all four good responses were 10.48% girls, and 44.44% boys. The difference is quite significant.

7.25% girls and 4.1% boys marked the bad answer (Figure 16).

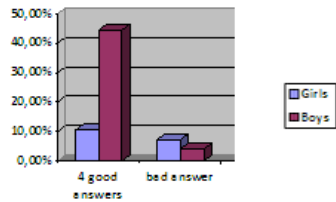


Figure 16

7. THE MAIN PROBLEMS AND CAUSES ACCORDING TO THE TEST

Students are often inattentive, uninterested. (They do not take their tasks seriously.)

Their thinking is not complex enough.

Based on drawing they are unable to conceive space forms.

They are not aware of the spatial position of the space elements – for example two lines are parallel or evasive, or perpendicular but evasive.

They cannot remember the name of the types of triangles.

They do not recognize the two-dimensional rectangles. They do not know the grouping of rectangles; namely, that the parallelogram is a rectangle and the rectangle and the parallelogram are trapezoids.

The spatial vision of most of students' is incomplete, weak or it is not advanced enough.

Turning the forms in their mind can cause problems as well.

Students do not think about all opportunities. They are not looking for more solutions.

8. ABOUT MY PLANS AND IDEAS

"According to Bishop (1980) and Ben-Chaim (1988) we can improve the spatial ability with training." [4]

I wrote a note according to the New Syllabus from Technical Description. Which contains the theoretical and practical elements. This can be divided into lessons, and can be downloaded and printed from my website (www.ritaart.hu).

In class 9.C., where we had only 16 lessons in the first semester, we often used my notes, especially the task sheets.

In other classes (9.D. and 13.C.) where we had more lessons, students did not follow the note.

We watched a lot of paintings from the Art History and analyzed the space delineations, and tried to apply their methods.

We could try some computer programs to modelling of the objects – but there are not enough computers, and time to experiment with

this way.

(It would be very useful and interesting to try some new innovations, for example 3D pen or 3D printer [5].)

9. CONCLUSIONS

It is very interesting that students who have already studied Technical Description are able to draw the objects in the way they see it but those who have not will use the usual scheme.

The students know quite a lot but unable to use them. The knowledge of the students is highly context-dependent they only know what was featured in the classroom during the lessons.

"People who have not heard of Descriptive Geometry do not even miss it, they can live without it [6]. Textbooks used to be full of geometrical mistakes concerning the figures (Figure 17: Orbit of the earth around the sun [7]; and Figure 18: The image of the earth). Luckily new textbooks and visuals are much better. Despite this fact the students draw the globe this way (Figure 18).

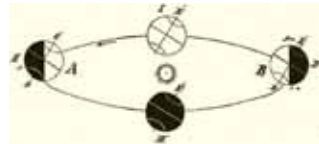


Figure 17

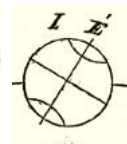


Figure 18

In geometry education needs some kind of reforms and change is needed, but unfortunately I have no idea, how to make my classes more effective. I am sure, the decline in the number of lessons causes the detriment of efficiency, especially if freehand drawing will disappear from the secondary school education.

Because of the changes currently taking place (New Syllabus and New Curriculum), many problems will appear only in the next school year.

I think, it is our mission to show different delineations in History of Art and in Descriptive Geometry to our students. Furthermore, we have to teach the students to wonder at the

world, and improve their space-vision and space representation. We should teach them carefully, let their experience develop in natural way.

ACKNOWLEDGMENTS

Thanks for Prof. Dr. József Szabó, University of Debrecen, my students from Belvárosi I. István Secondary School.

REFERENCES

- [1] <http://www.mmo.njszt.hu/Kiadvanyok/2002/cikkek/kosztyan.pdf>
- [2] cf.: http://ganymedes.lib.unideb.hu:8080/dea/bitstream/2437/163373/1/katona_janos_disszertacio.pdf
- [3] http://www.sulinet.hu/tanar/kompetenciateruletek/2_matematika/4_modszertani_segedletek/3_hattertan-felsotag-kozepiskola/terszemleletfejl_te_is_latod_i.pdf
- [4] Mária Bakó. Different projecting methods in teaching spatial geometry. http://www.erne.tu-dortmund.de/~erne/CERME3/Groups/TG7/TG7_Bako_cerme3.pdf
- [5] LIX Is The World's Smallest 3D Drawing Pen That Lets You Draw In The Air, <http://www.boredpanda.com/lix-3d-printing-pen/>
- [6] József Szabó. Ábrázoló geometria tanárszak története Debrecenben 1956-2011; Debrecen, 2013, magánkiadás; ISBN:978-963-08-6646-0
- [7] <http://mek.oszk.hu/04700/04759/html/madarehegy0016.html>

ABOUT THE AUTHOR

1. Rita Kiss-György, teaches Art and Design and Technical Description in some classes at Belvárosi I. István Secondary School in Székesfehérvár, and she is a painter and applied graphic artist. She has a degree in Art, an MSc in Mathematics and an MSc in Descriptive Geometry (in Debrecen). Her research interests are Descriptive Geometry in Art History, Reversed Perspective in the

Icons and space representation in Art lessons. She can be reached by e-mail: kgrita@freemail.hu, or www.ritaart.hu.

ON THE SPECIAL SURFACES THROUGH THE ABSOLUTE CONIC WITH A SINGULAR POINT OF THE HIGHEST ORDER

Sonja GORJANC and Ema JURKIN
University of Zagreb, Croatia

ABSTRACT: In this paper we observe a special class of surfaces in the Euclidean space E^3 which touch the plane at infinity through the absolute conic and have a singular point of the highest order. We study their properties and visualize them with the program *Mathematica*.

Keywords: surfaces, absolute conic, tangent cone, singular point

1. INTRODUCTION

In this paper we will study one special class of surfaces in the Euclidean space E^3 so let us start by recalling some definitions and facts about the surfaces. In the real three-dimensional projective space $P^3(\mathbb{R})$, in homogeneous Cartesian coordinates $(x:y:z:w)$, $(x,y,z \in \mathbb{R}, w \in \{0,1\})$, $(x:y:z:w) \neq (0:0:0:0)$, the equation

$$F_n(x,y,z,w) = 0,$$

where F_n is a homogeneous algebraic polynomial of degree n , defines an n^{th} order surface S_n . This equation can be written as

$$f_n(x,y,z) + wf_{n-1}(x,y,z) + \dots + w^{n-1}f_1(x,y,z) + w^n f_0(x,y,z) = 0,$$

where f_i , $i = 0, 1, \dots, n$, are homogeneous algebraic polynomials of degree i .

Any straight line, not lying on S_n , intersects S_n in n points and any plane intersects S_n in the n^{th} order plane curve.

A point T of the surface S_n for which at least one partial derivation of F_n is not equal to zero is called the regular point of S_n . All tangents to the surface at that point lie in one plane - the tangent plane of S_n at T .

A point T of the surface S_n for which all partial derivations of F_n are equal to zero is called the singular point of S_n . The tangents to S_n at this point form an algebraic cone with vertex in

T . If the tangent cone is of order k , the point T is the k -fold point of the surface S_n . Every plane through T intersects S_n in the n^{th} order plane curve with the k -fold point in T .

If the origin $O(0:0:0:1)$ is the k -fold point of S_n , then S_n has the equation

$$f_n(x,y,z) + wf_{n-1}(x,y,z) + \dots + w^{n-k}f_k(x,y,z) = 0, \quad (1)$$

and the tangent cone at O is given by

$$f_k(x,y,z) = 0. \quad (2)$$

In the paper [2] the author studied the quartics Φ_4 in E^3 which have a triple point and touch the plane at infinity through the absolute conic. The surfaces were classified according to the type of the tangent cone \mathcal{T} at the triple point. The following cases were observed: \mathcal{T} is a proper 3-order cone, \mathcal{T} splits into a proper 2-order cone and a real plane, \mathcal{T} splits into three planes. In this paper we give the generalization of the third class of surfaces Φ_4 .

2. SURFACES \mathcal{S}_{2n} WITH A $(2n-1)$ -FOLD POINT TOUCHING THE PLANE AT INFINITY THROUGH THE ABSOLUTE CONIC

In the real projective space $P^3(\mathbb{R})$ the Euclidean metric defines the Euclidean space E^3 with the absolute conic given by the equations: $x^2 + y^2 + z^2 = 0, w = 0$.

Theorem 1 A surface \mathcal{S}_{2n} given by the equation

$$A_2(x, y, z)^n + wH_{2n-1}(x, y, z) = 0, \quad (3)$$

where $A_2(x, y, z) = x^2 + y^2 + z^2$ and $H_{2n-1}(x, y, z)$ is a product of $2n - 1$ linear homogeneous polynomials, is a surface of order $2n$ which has a $(2n - 1)$ -fold point in the origin and intersects the plane at infinity only in the absolute conic.

Proof. By comparing (3) with (1) it is evidently that the origin $O(0:0:0:1)$ is the $(2n - 1)$ -fold point of \mathcal{S}_{2n} at which the tangent cone splits into $2n - 1$ planes. The intersection of the plane at infinity ($w = 0$) and the surface \mathcal{S}_{2n} is given by $A_2(x, y, z)^n = 0$. It is the absolute conic with the intersection multiplicity n . \square

Theorem 2 There are only $2(2n - 1)$ straight lines through the origin lying entirely on the surface \mathcal{S}_{2n} . They are the intersections of cones given by $A_2(x, y, z)^n = 0$ and $H_{2n-1}(x, y, z) = 0$. They are imaginary in pairs.

Proof. Let a line p through $O(0:0:0:1)$ be spanned by O and a further point $P(a:b:c:1) \neq O$. The line p is parametrized by

$$p \quad \dots \quad (x:y:z:1) = (at:bt:ct:1), t \in \mathbb{R}.$$

It lies on \mathcal{S}_{2n} if and only if

$$A_2(at, bt, ct)^n + H_{2n-1}(at, bt, ct) = 0,$$

for every $t \in \mathbb{R}$. This is precisely when

$$t^{2n-1}[tA_2(a, b, c)^n + H_{2n-1}(a, b, c)] = 0,$$

for every $t \in \mathbb{R}$. It follows that $A_2(a, b, c)^n = 0$, $H_{2n-1}(a, b, c) = 0$. Therefore, $A_2(at, bt, ct)^n = 0$, $H_{2n-1}(at, bt, ct) = 0$, for every $t \in \mathbb{R}$. Evidently the line p lies on the cones given by equations $A_2(x, y, z)^n = 0$ and $H_{2n-1}(x, y, z) = 0$. We conclude: the only lines through the origin that lie on \mathcal{S}_{2n} are the isotropic lines in the tangent planes at the origin. \square

Theorem 3 The surface \mathcal{S}_{2n} has only one real singular point.

Proof. Let us suppose that there is a k -fold point $T \neq O$ of \mathcal{S}_{2n} , $k \geq 2$. The real line OT intersects \mathcal{S}_{2n} in the point O with the intersection multiplicity $2n - 1$ and the point T with the intersection multiplicity k , or entirely lies on \mathcal{S}_{2n} . The first option is not possible because the order of the \mathcal{S}_{2n} is $2n < 2n - 1 + k$, while the second option is in contradiction with Theorem 2. \square

Theorem 4 The surface \mathcal{S}_{2n} touches the plane at infinity through the absolute conic.

Proof. In Theorem 1 it stated that the absolute conic is the intersection of \mathcal{S}_{2n} and the plane at infinity with the intersection multiplicity n . It is left to show that the absolute conic is not the singular line of \mathcal{S}_{2n} . If the absolute conic was the singular line of \mathcal{S}_{2n} , its every point would be the singular point of \mathcal{S}_{2n} and every isotropic line through the origin O would lie on the surface \mathcal{S}_{2n} . This is not possible since there are only $2(2n - 1)$ straight lines through the origin lying entirely on \mathcal{S}_{2n} . \square

According to Theorem 3 there is no real double point on the surface \mathcal{S}_{2n} and therefore there is no selfintersections of \mathcal{S}_{2n} . Hence, the surface consists of the separated parts sharing only $(2n - 1)$ -fold point O . These parts we will call *petals*.

Theorem 5 The maximum number of petals of the surface \mathcal{S}_{2n} in the $(2n - 1)$ -fold point equals $2n^2 - 3n + 2$.

Proof. Each petal of \mathcal{S}_{2n} lies on one side of one tangent plane at the $(2n - 1)$ -fold point $O(0:0:0:1)$. Therefore, the number of petals is twice less than the number of parts into which space is divided by $2n-1$ planes passing through one point. Let us first show that the maximum number of parts of space divided by k copunctal planes equals $k^2 - k + 2$. The number of the parts will be the largest when no three planes contain a common line. We will assume that this condition is fulfilled. The plane is divided by k concurrent

lines into $2k$ parts. Let the number of parts of the space divided by k plane be denoted by \bar{k} . The additional $(k+1)^{st}$ plane intersects the first k planes into k lines which divide the plane into $2k$ regions. Therefore, $\bar{k} + 1 = \bar{k} + 2k$. We have

$$\begin{aligned}\bar{1} &= 2 \\ \bar{2} &= \bar{1} + 2 \\ &\vdots \\ \bar{k} &= \bar{k-1} + 2(k-1).\end{aligned}\quad (4)$$

By summation of these equations we obtain the following: $\bar{k} = 2 + 2 + 4 + \dots + 2(k-1) = 2 + 2 \cdot \frac{(k-1)k}{2} = k^2 - k + 2$. By substituting k with $2n-1$, we get $\bar{2n-1} = 4n^2 - 6n + 4$ and obtain the claimed result. \square

Remark. If no three tangent planes at the origin share a common line, the number of petals equals $2n^2 - 3n + 2$. If three planes intersect at a line, the number of petals decreases. Let us prove that if l planes pass through the same line, the number of petals is decreased by $\frac{(l-1)(l-2)}{2}$. Let us first take into consideration $k-l$ planes such that no three planes share a common line. Then we add two of l planes with a common line, and at the end we add remaining $k-l-2$ planes. The list of equations (4) now becomes

$$\begin{aligned}\bar{1} &= 2 \\ \bar{2} &= \bar{1} + 2 \\ &\vdots \\ \bar{k-l} &= \bar{k-l-1} + 2(k-l-1) \\ \bar{k-l+1} &= \bar{k-l} + 2(k-l) \\ \bar{k-l+2} &= \bar{k-l+1} + 2(k-l+1) \\ \bar{k-l+3} &= \bar{k-l+2} + 2(k-l+1) \\ &= \bar{k-l+2} + 2(k-l+2) - 2 \\ \bar{k-l+4} &= \bar{k-l+3} + 2(k-l+1) \\ &= \bar{k-l+3} + 2(k-l+3) - 4 \\ &\vdots \\ \bar{k-1} &= \bar{k-2} + 2(k-l+1) \\ &= \bar{k-2} + 2(k-2) - 2(l-3)\end{aligned}$$

$$\begin{aligned}\bar{k} &= \bar{k-1} + 2(k-l+1) \\ &= \bar{k-1} + 2(k-1) - 2(l-2).\end{aligned}$$

Therefore, $\bar{k} = 2 + 2 \cdot \frac{(k-1)k}{2} - 2 \cdot \frac{(l-2)(l-1)}{2} = k^2 - k + 2 - (l-1)(l-2)$. Since the number of the petals of \mathcal{S}_{2n} is two times less then the number of the parts into which space is divided by the tangent planes at the origin, we can conclude the following: if l of $2n-1$ tangent planes intersect in one line, the number of the petals is decreased by $\frac{(l-1)(l-2)}{2}$.

2.1 Parametric equations of surfaces \mathcal{S}_{2n}

By substituting $\omega = 1$ into the equation (3) we get the following equation of the surface \mathcal{S}_{2n} :

$$A_2(x, y, z)^n + H_{2n-1}(x, y, z) = 0. \quad (5)$$

If we use the spherical coordinates (ρ, ϕ, θ) :

$$x = \rho \cos \phi \sin \theta, \quad y = \rho \sin \phi \sin \theta, \quad z = \rho \cos \theta,$$

the equation (5) takes the following form:

$$\rho^{2n-1} (\rho + H_{2n-1}(\cos \phi \sin \theta, \sin \phi \sin \theta, \cos \theta)) = 0.$$

For every point of of the surface \mathcal{S}_{2n} , except for the $(2n-1)$ -fold point $O(0, 0, 0)$, it holds

$$\rho = -H_{2n-1}(\cos \phi \sin \theta, \sin \phi \sin \theta, \cos \theta),$$

and therefore the surface \mathcal{S}_{2n} is given by the parametric equations:

$$\begin{aligned}x(u, v) &= -H_{2n-1}(\cos \phi \sin \theta, \sin \phi \sin \theta, \cos \theta) \cos \phi \sin \theta, \\ y(u, v) &= -H_{2n-1}(\cos \phi \sin \theta, \sin \phi \sin \theta, \cos \theta) \sin \phi \sin \theta, \\ z(u, v) &= -H_{2n-1}(\cos \phi \sin \theta, \sin \phi \sin \theta, \cos \theta) \cos \theta, \\ \phi, \theta &\in [0, \pi] \times [0, \pi].\end{aligned}\quad (6)$$

$H_{2n-1}(x, y, z)$ is the product of $2n-1$ linear polynomials in x, y, z and therefore determined by $3(2n-1)$ coefficients.

2.2 Visualization of surfaces \mathcal{S}_{2n}

Based on the equations (5) or (6), we can visualize any surface \mathcal{S}_{2n} with the program *Mathematica*.



Figure 1: Three surfaces \mathcal{S}_n , where $n = 2, 3, 4$, are shown in this figure. The tangent cone at the origin splits into $2n - 1$ planes that intersect through the axis z . Each tangent cone has $4n - 2$ planes of symmetry, and corresponding surface \mathcal{S}_n has $2n - 1$ petals and $2n - 1$ planes of symmetry.

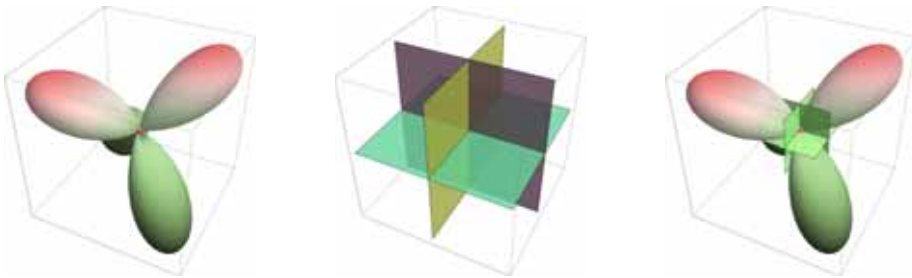


Figure 2: This figure shows one surface \mathcal{S}_4 and its tangent cone at the origin that splits into three coordinate planes. The surface is given by the following implicit equation: $A_2(x, y, z)^2 + xyz = 0$. Since three tangent planes at the origin have only one common point, the surface has 4 petals.

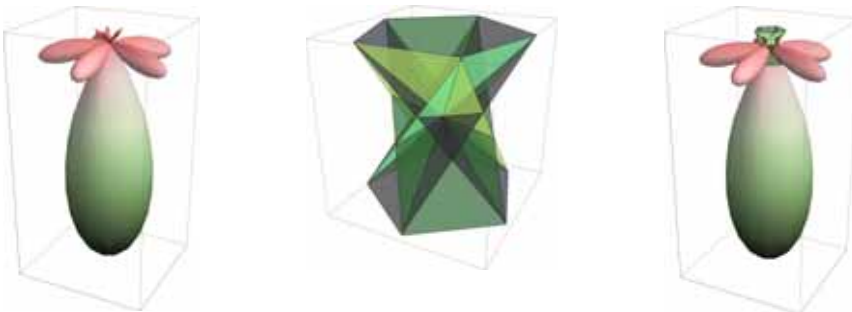


Figure 3: This figure shows one surface \mathcal{S}_6 and its tangent cone at the origin that splits into 5 planes which have only one common points. The surface has the largest number of petals, i.e. 11 petals.



Figure 4: This figure shows one surface \mathcal{S}_8 and its tangent cone at the origin that splits into 5 planes, where 4 planes share the common axis z . Therefore, the largest number of petals (11) decreases to 8.



Figure 5: This figure shows one surface \mathcal{S}_8 and its tangent cone at the origin that splits into 7 planes and there exist 6 lines which are the intersections of 3 tangent planes. Therefore, the largest number of petals (22) decreases to 16.

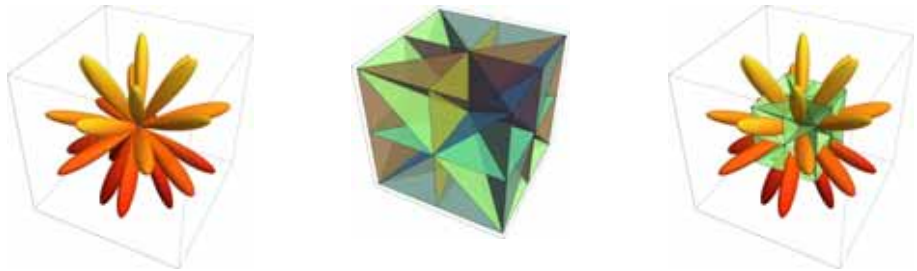


Figure 6: This figure shows one surface \mathcal{S}_{10} and its tangent cone at the origin that splits into 9 planes. There exist 3 lines that are the intersections of 4 tangent planes, and 4 lines which are the intersections of 3 tangent planes. Therefore, the largest number of petals (37) decreases to 24.

3. CONCLUSIONS

In this paper we observe a special class of surfaces \mathcal{S}_{2n} in the Euclidean space, given by the equation of the form

$$A_2(x, y, z)^n + wH_{2n-1}(x, y, z) = 0,$$

where $A_2(x, y, z) = x^2 + y^2 + z^2$ and $H_{2n-1}(x, y, z)$ is a product of $2n - 1$ linear homogeneous polynomials. We show that \mathcal{S}_{2n} is a surface of order $2n$ which has a $(2n - 1)$ -fold point in the origin and touches the plane at infinity through the absolute conic. The surface consists of the separated parts (petals) sharing only one real $(2n - 1)$ -fold point. We prove that the largest number of petals of \mathcal{S}_{2n} equals $2n^2 - 3n + 2$ and show how this number decreases if some tangent planes at the origin pass through same straight line.

REFERENCES

- [1] K. Fladt and A. Baur. *Analytische Geometrie spezieller Flächen und Raumkurven*. Friedr. Vieweg & Sohn, Braunschweig, 1975.
- [2] S. Gorjanc. Quartics in E^3 which Have a Triple Point and Touch the Plane at Infinity through the Absolute Conic. *Mathematical Communications*, 9: 67–78, 2004.
- [3] S. Gorjanc. Rose Surfaces and their Visualizations. *Journal for Geometry and Graphics*, 13(1): 59–67, 2010.
- [4] J. Harris. *Algebraic Geometry*. Springer, New York, 1995.
- [5] G. Salmon. *A Treatise on the Analytic Geometry of Three Dimensions*. Vol.I., Chelsea Publishing Company, New York, (reprint), 1958.
- [6] G. Salmon. *Higher Plane Curves*. Chelsea Publishing Company, New York, (reprint), 1960.

ABOUT THE AUTHORS

1. Sonja Gorjanc, Ph. D., is Assistant Professor at the Faculty of Civil Engineering (Department of Mathematics) at the University of Zagreb, the editor-in-chief of the scientific-professional journal *KoG* and the former president of the *Croatian Society for Geometry and Graphics*. Her research interests are in projective and Euclidean geometry, *Mathematica* computer graphics, curricular developments in geometry and graphics. She can be reached by e-mail: sgorjanc@grad.hr.
2. Ema Jurkin, Ph. D., is Assistant Professor at the Faculty of Mining, Geology and Petroleum Engineering (Department of Mathematics, Informatics and Descriptive Geometry) at the University of Zagreb, an editor of the scientific-professional journal *KoG* and the vice-president of the *Croatian Society for Geometry and Graphics* and the vice-president of *International Society for Geometry and Graphics* for the region Europe\Near East\Africa. Her research interests are projective geometry, Euclidean and non-Euclidean planes, treated by synthetic (constructive) and analytical methods. She can be reached by e-mail: ema.jurkin@rgn.hr.

STUDENTS' ABILITY TO MODEL PARTS FROM VARIOUS TYPES OF DRAWINGS AND THE RELATIONSHIP TO OTHER MEASURES IN ENGINEERING GRAPHICS COURSES

Theodore BRANOFF¹ and Modris DOBELIS²

¹Illinois State University, U.S.A. ²Riga Technical University, Latvia

ABSTRACT: During the Fall 2013 and Spring 2014 semesters, a study was conducted where students completed the PSVT:R and the MCT and then were asked to complete three different modeling tasks. These tasks included modeling a part when given the object in the context of an assembly drawing (one part within the original modeling test), modeling a part when given an isometric pictorial of the object, and finally modeling a part when given a detail drawing of the object. Research questions for this study include the following: When evaluating 3D models within the modeling test, are models evaluated consistently when using a new rubric and the original rubric? Is there a relationship between a student's spatial visualization ability and their ability to model a part from a pictorial, assembly drawing or detail drawing? Is the MCT still a better predictor of a student's modeling success than the PSVT:R? This study will present the results from these research questions, examine how students approached modeling parts, give suggestions for how instruction in the course can be improved, as well as give recommendations for future studies.

Keywords: engineering graphics literacy, constraint-based CAD, spatial visualization, modeling strategies.

1. INTRODUCTION

A recent series of studies investigated students' ability to model parts when given assembly drawing information [4–7, 12]. These investigations revealed that the developed modeling test, as measured by the original rubric [6], had mixed results when examining relationships with measures in the course (e.g., final project and final exam). Some of these studies also examined whether students' modeling ability was related to their spatial visualization ability [4–5, 7]. There were positive correlations between the PSVT:R and the modeling test (not all were significant) and significant positive correlations between the MCT and the modeling test. Recommendations included repeating the study using a shorter modeling activity, examining a more efficient way of evaluating the models, and using qualitative methods for evaluating modeling strategies.

2. REVIEW OF LITERATURE

As educators, what do we look for when we evaluate students' 3D CAD models? Do we have a set list of items that each student must demonstrate? Should the model exhibit certain design intent when modified? What does it mean to be an expert in the area of constraint-based CAD modeling? Educators have been studying the modeling strategies of students and practitioners for approximately 15 years. The major topics of these studies involve CAD modeling strategies, conceptual framework research, and methods of evaluating models.

Expertise in CAD requires more than just instruction and experience using the software. Experts have a background about their products, know good modeling strategies, and know when to apply the modeling strategies [8]. New

CAD users who only have a background in engineering graphics and descriptive geometry still struggle to build appropriate design intent into their models [10].

What are the important themes that experts in constraint-based modeling exhibit? After studying practicing professionals, Hartman identified several core themes that are typically demonstrated by experts [14-16]. These themes include strategy for tool use, problem definition and solution, design considerations, domain knowledge, and professional and academic experiences. He also noted the subordinate themes of software techniques, downstream uses of the CAD model, technical communication, social communication, requisite CAD model characteristics, and problem solving techniques. Finally, he talks about the transitional themes of design environment, the way the expert worked, support structures, artifacts, personal characteristics, typical domain activities, conceptions of expertise, and factors related to CAD usage. Rynne, Gaughran, and Seery define the components of CAD expertise as the part modeling task, procedural 3D CAD knowledge, strategic 3D CAD knowledge, declarative 3D CAD knowledge, graphical visualization capability, modeling deconstruction capability, and metacognitive processes [18].

There has been a wide range of studies that have examined evaluating CAD models. Some of these have involved automatic evaluation of the models through electronic means, and others have involved evaluating the models by opening up each student's file. In each case it is important that the instructor identify the key components that they want the student to exhibit in their model. Steinhauer evaluates students' models based on the general categories of approach, structure, accuracy, robustness, and creativity [19]. Ault and Fraser look for correct geometry, appropriate choice and order of features, proper location of origin, proper view orientation, use diameter and radius dimensions correctly, correct hole placement, use of reference geometry for dependent features, and general modeling strategy [1]. Baxter and

Guerri developed an automated online evaluation system. Their system allowed each instructor to determine how the student's model was queried. For example, an instructor may require the first feature of a model to be a revolve instead of an extrude. If the student used an extrude they would not receive any points. Students could submit their files as many times as necessary before the assignment deadline [2, 3]. Company, Contero & Salvado-Herranz explain that quality CAD models are *valid* if they can be opened by suitable applications without errors, are *complete* if they include all product aspects relevant for design applications, are *consistent* if they do not exhibit errors when modified, a *concise* when they do not include irrelevant information or features, and are *effective* when they convey correct design intent [9].

More recently, studies have been conducted to examine students' ability to build specific design intent into models. Devine & Laingen outline a procedure used in their course that students can use to self-assess their model [11]. Students are given two self-check opportunities where they must measure a distance, one face area, and the total face area of the part. For the second self-check, students are required to change several dimensions on the part before measuring. This allows the instructor to determine if the initial dimensions captured the correct design intent. Peng et al. advocate for a similar approach [17].

3. RESEARCH QUESTIONS

The current study was designed to conduct an investigation into using an alternative rubric for assessing models. In addition, students were given alternative types of drawings (a pictorial and a detail drawing) to model, which were evaluated using the new rubric. To continue the theme of the previous engineering graphics literacy studies, correlations between the scores on the modeling activities and two standard visualization tests were also examined. The research questions for this study were:

1. Is there a difference in score between a model evaluated by the original rubric [6] and the same model evaluated by the new rubric?
2. Is there a relationship between a student's spatial visualization ability and their ability to model a part from an assembly drawing?
3. Is there a relationship between a student's spatial visualization ability and their ability to model a part from a pictorial drawing?
4. Is there a relationship between a student's spatial visualization ability and their ability to model a part from a detail drawing?
5. As with the previous studies, is the Mental Cutting Test a better predictor of a student's modeling ability than the Purdue Spatial Visualization Test: Visualization of Rotations?

4. PARTICIPANTS

The participants for this study were enrolled in a junior-level constraint-based modeling course at North Carolina State University during the fall 2013 and spring 2014 semesters. The course consists of engineering graphics standards and conventional practices (sectional views, dimensioning, threads & fasteners, and working drawings), geometric dimensioning and tolerancing, and constraint-based modeling techniques (assemblies, advanced drawing applications, macros, design tables, and rendering). Tables 1-3 summarize the demographic information of the participants.

Table 1: Gender of Participants.

Gender	Freq.	%
Female	4	9.52%
Male	38	90.48%
TOTAL	42	100.00%

Table 2: Academic Year of Participants.

Year	Freq.	%
Sophomore	5	11.90%
Junior	21	50.00%
Senior	16	38.10%
TOTAL	42	100.00%

Table 3: Academic Major of Participants.

Major	Freq.	%
Agricultural & Environmental Technology	2	4.76%
Biological Engineering	1	2.38%
Civil Engineering	2	4.76%
Computer Science	1	2.38%
Electrical Engineering	1	2.38%
Mechanical & Aerospace Engineering	17	40.48%
Technology, Engineering, & Design Education	18	42.86%
TOTAL	42	100.00%

Most of the students enrolled in the course were male students in their third or fourth year from either engineering majors or technology, engineering & design education. The technology, engineering & design education students take the course as part of their major requirements. Other students typically take the course as part of a 15 credit hour minor in graphic communications.

5. MODELING ACTIVITIES

Three different modeling activities were used in this study. Figure 1 shows the assembly drawing that was used in previous studies. In those studies, students were asked to model as many of the 7 parts as possible in a 110 minute class period. For the current study, students were only asked to model one of the parts in the given assembly – the SET SCREW (Figure 2). As with the previous studies, students were given a metric ruler and asked to scale dimensions directly off of the drawing. The assembly drawing was created with a drawing scale of 2:1. Only overall dimensions and a few other dimensions required for installation were given, including thread designations and sizes.

In addition to the SET SCREW, students were asked to model the INDEX ARM (Figure 3) and the RING (Figure 4). The INDEX ARM was selected because of its mixture of extrudes, cuts, holes, fillets, and symmetry. It has also been used in at least one other study at another

institution, so some general comparisons might be made [19]. The RING was chosen for the detail drawing exercise because it is similar to existing parts in industry. It also requires students to be able to determine best strategies for using revolves and circular patterns.

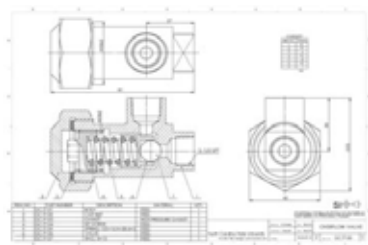


Figure 1: Modeling Test Assembly Drawing.



Figure 2: Model of SET SCREW.

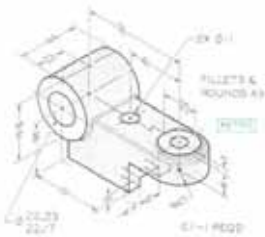


Figure 3: INDEX ARM [13].

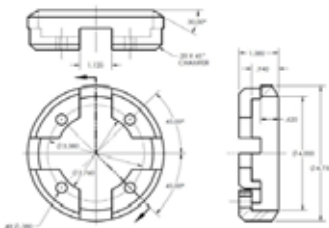


Figure 4: RING Drawing.

6. METHODOLOGY

During the latter part of the fall 2013 and spring 2014 semesters, students were given a lecture on reading international drawing standards. They were then given a practice problem similar to the assembly drawing in Figure 1. The instructor guided students through modeling several of the parts in the assembly. Within one week of this exercise students also completed the PSVT:R and the MRT spatial tests. Both tests were administered electronically through the learning management system for the course. During the next class students were asked to model the SET SCREW, INDEX ARM, and the RING. They were given 30 minutes to complete each model. All models were saved to a centralized server, and the researcher copied all of the files to a local computer for evaluation.

Two rubrics were used in the assessment of the models. The original assessment rubric spreadsheet [6] was created to account for model accuracy and time required to model each part. Each feature and sketch (if any) was analyzed individually. Penalty points were assigned for each wrong geometric dimension including under-defined sketches. Penalty points were added for each dimension of the geometric primitive missing in the model, incorrect dimensions, and failure to correctly represent cosmetic threads. A new rubric (Figure 5) was developed based on themes gathered from the review of literature. Although some of the researchers applied point values to each category in their rubrics, there was no consistency in the weighting of the categories. For the current study, weights were given on the basis of the relevant importance of each category to the problem solution.

Category	Possible Points	Points Awarded
Base/Core feature correctly identified	10	
Orientation of initial sketch plane	10	
Best model origin	10	
Sketches are simple, fully constrained, and reflect appropriate design intent	20	
Appropriate feature end-conditions	10	
Correct application of symmetry/duplication	10	
Accuracy/Complete Model	20	
Modeling strategy efficiency	10	
TOTAL	100	

Figure 5: New Rubric to Evaluate Models.

The models were evaluated in the following order:

1. All SET SCREW models were evaluated with the original rubric.
2. All SET SCREW models were evaluated with the new rubric.
3. All INDEX ARM models were evaluated with the new rubric.
4. All RING models were evaluated with the new rubric.

7. ANALYSIS OF RESULTS

Table 4 displays the descriptive statistics for the study. All students completed the two visualization tests and made some attempt at the SET SCREW, and the RING. One student did not complete any part of the INDEX ARM.

Table 4: Descriptive Statistics.

Activity	N	Min	Max	Mean	SD
SET SCREW †	42	0.00	95.00	63.33	26.31
SET SCREW *	42	0.00	99.00	80.29	20.85
INDEX ARM *	42	0.00	100.00	78.83	17.78
RING *	42	23.00	98.00	73.95	18.75
PSVT:R	42	10.00	30.00	24.74	5.32
MCT	42	2.00	25.00	16.88	5.64

† Evaluated with Original Rubric

* Evaluated with New Rubric

Research Question 1: Is there a difference between a model evaluated by the original rubric [6] and the same model evaluated by the new rubric? Figure 6 displays the scores for each student's SET SCREW model evaluated

by the original rubric and the new rubric. The graph reveals large differences in scores for models that were low in quality. There is less discrepancy in the two rubrics when the models are closer to the correct solution. Table 6 displays the results of a paired-sample t-test to evaluate whether there is a difference between the two rubrics.

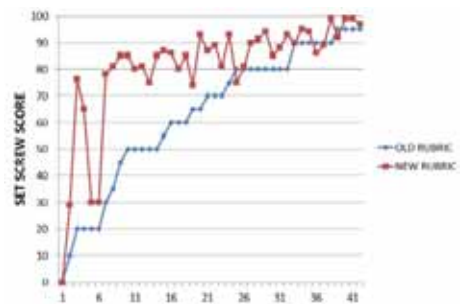


Figure 6: SET SCREW Scores Evaluated by Original Rubric and New Rubric.

Table 6: Paired-Sample t-test.

Paired Diff			t	df	Sig
Mean	SD	Std Error			
-16.95	15.76	2.43	-6.97	41	.000 *

* Significant at $\alpha=.05$ level.

The analysis revealed a significant difference in SET SCREW scores when using the two rubrics. Scores tended to be higher when evaluating the models using the newer rubric.

Research Questions 2-4: Is there a relationship between a student's spatial visualization ability and their ability to model a part from an assembly drawing, pictorial drawing, or detail drawing?

Research Question 5: Is the Mental Cutting Test a better predictor of a student's modeling ability than the Purdue Spatial Visualization Test: Visualization of Rotations?

Scatterplots were created to examine the visual relationships between the scores on each modeling activity and scores on the two spatial visualization tests. These scatterplots are displayed in Figures 7-15.

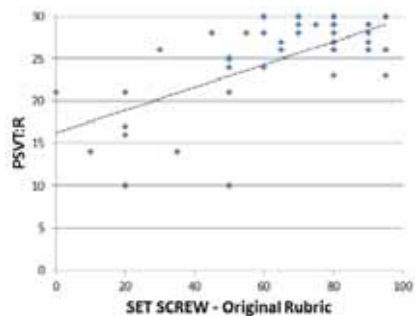


Figure 7: SET SCREW – Original Rubric vs. PSVT:R.

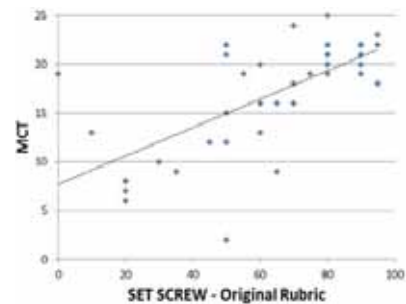


Figure 8: SET SCREW – Original Rubric vs. MCT.

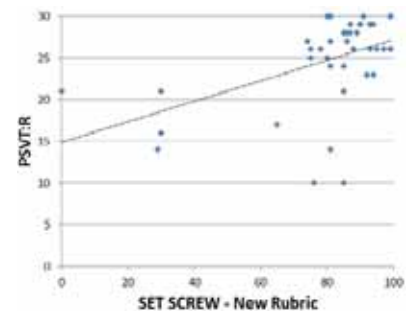


Figure 9: SET SCREW – Original New vs. PSVT:R.

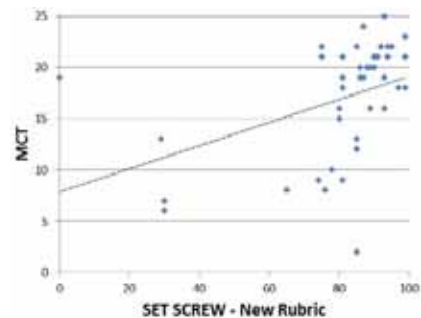


Figure 10: SET SCREW – Original New vs. MCT.

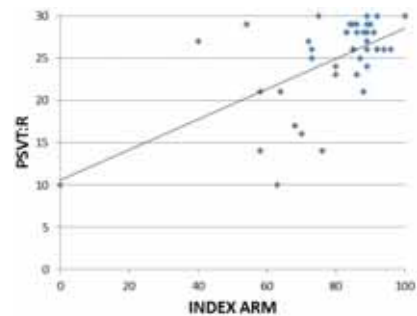


Figure 11: INDEX ARM vs. PSVT:R.

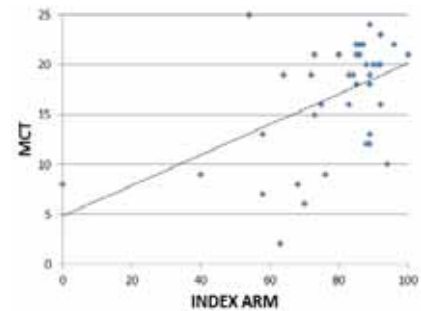


Figure 12: INDEX ARM vs. MCT.

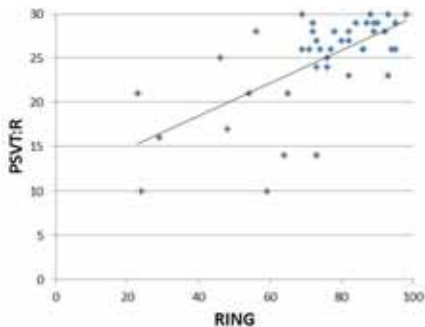


Figure 13: RING vs. PSVT:R.

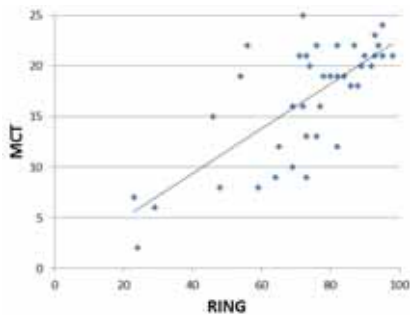


Figure 14: RING vs. MCT.

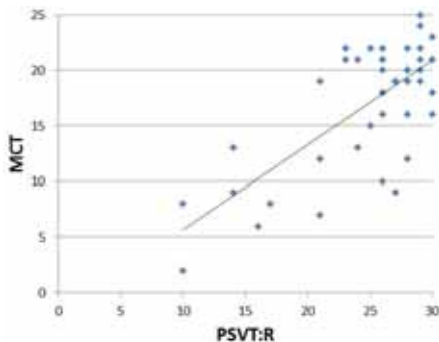


Figure 15: PSVT:R vs. MCT.

The scatterplots display positive relationships between the modeling activities and the two spatial visualization tests. The descriptive statistics show that some of the scores were very spread out with large standard deviations. Since

the distributions of the data do not appear to be normal, a non-parametric Spearman's Rho was used to test research questions 2-5. Table 7 displays the results of this analysis.

Table 7: Spearman's Rho Correlation between the PSVT:R and the Modeling Test.

Spearman's Rho		SET SCREW Original Rubric	SET SCREW New Rubric	INDEX ARM	RING	PSVT:R	MCT
SET SCREW Original Rubric	Corr Coef.	1.000					
	Sig (2 tailed)						
	N	42					
SET SCREW New Rubric	Corr Coef.	.817**	1.000				
	Sig (2 tailed)	.000					
	N	42	42				
INDEX ARM	Corr Coef.	.438**	.518**	1.000			
	Sig (2 tailed)	.004	.000				
	N	42	42	42			
RING	Corr Coef.	.700**	.726**	.619**	1.000		
	Sig (2 tailed)	.000	.000	.000			
	N	42	42	42	42		
PSVT:R	Corr Coef.	.533**	.457**	.449**	.589**	1.000	
	Sig (2 tailed)	.000	.002	.003	.000		
	N	42	42	42	42	42	
MCT	Corr Coef.	.662**	.554**	.373*	.627**	.523**	1.000
	Sig (2 tailed)	.000	.000	.015	.000	.000	
	N	42	42	42	42	42	42

** . Correlation is significant at the 0.01 level (2-tailed).

* . Correlation is significant at the 0.05 level (2-tailed).

The Spearman's Rho analyses show significant positive correlations between scores on all of the modeling activities and both visualization tests. As one might expect, there was also a significant positive correlation between the PSVT:R and the MCT ($\rho = .523$, $\alpha = .000$).

8. CONCLUSIONS

This study revealed a significant difference in scores on the SET SCREW when the parts were evaluated with two different rubrics.

Scores were significantly higher when evaluated with the new rubric than with the original rubric. The original rubric tended to penalize students more for making errors within each feature, which resulted in much lower scores. Part of this is the result of having different expectations of a model within each rubric. The purpose of the rubric is to clearly define what one should look for when evaluating the model. Since the scores were different when the two rubrics were used, it is obvious that they each define different parameters. If the goal of the rubric is just to assess a final examination of the student's work, each of these rubrics may be appropriate. If educators are more interested in teaching good modeling strategies, than we should strive to develop instructional methods that help students understand appropriate design intent for a specific situation by giving them an opportunity to change the parameters within a model. Some of these instructional practices are already outlined in previous research [12, 18, 22].

The Spearman Rho analyses of the data revealed that there are significant correlation between students' scores on all of the modeling activities and the two spatial visualization tests. Students who scored higher on the PSVT:R and MCT tended to score higher on the modeling activities. Consistent with previous studies [5], the MCT had a higher correlation value with the modeling activities than the PSVT:R. The review of literature clearly outlines other factors that are involved in students' ability to model parts from detail drawings, assembly drawings, and pictorial drawings (e.g., problem definition and solution, design considerations, domain knowledge, and professional and academic experiences, strategic CAD knowledge, procedural CAD knowledge, etc.). The results of this study and others indicate that spatial visualization ability is an important factor in students' constraint-based modeling ability.

REFERENCES

- [1] Ault, H. K., & Fraser, A. (June, 2013). A comparison of manual vs. online grading for solid models. *Proceedings of the 2013 ASEE Annual Conference, Atlanta, Georgia, June 23-26, 2013.*
- [2] Baxter, D.H. (2003). Evaluating an automatic grading system for an introductory computer aided design course. *Proceedings of the 58th Annual Midyear Conference of the Engineering Design Graphics Division of the American Society for Engineering Education, Scottsdale, Arizona, November 16-19, 2003.*
- [3] Baxter, D.H. & Guerci, M. J. (2003). Automating an introductory computer aided design course to improve student evaluation. *Proceedings of the 2003 Annual Conference of the American Society for Engineering Education, Nashville, Tennessee, June 22-25, 2003.*
- [4] Branoff, T. J., & Dobelis, M. (October, 2013). Spatial visualization ability and students' ability to model objects from engineering assembly drawings. *Proceedings of the 68th Midyear Conference of the Engineering Design Graphics Division of the American Society for Engineering Education, Worcester, Massachusetts, October 20-22, 2013.*
- [5] Branoff, T. J., & Dobelis, M. (June, 2013). The relationship between students' ability to model objects from assembly drawing information and spatial visualization ability as measured by the PSVT:R and MCT. *Proceedings of the 2013 Annual Meeting of the American Society for Engineering Education, Atlanta, Georgia, June 23-26, 2013.*
- [6] Branoff, T. J., & Dobelis, M. (January, 2012). Engineering graphics literacy: Measuring students' ability to model objects from assembly drawing information. *Paper published in the proceedings of the 66th Midyear Conference of the Engineering Design Graphics Division of the Amer-*

- ican Society for Engineering Education, Galveston, Texas, January 22-24, 2012.
- [7] Branoff, T. J., Dobelis, M. (2012). The relationship between spatial visualization ability and students' ability to model 3D objects from engineering assembly drawings. *Engineering Design Graphics Journal*, 76 (3), 37-43.
- [8] Chester, I. (2008). 3D-CAD: Modern technology – outdated pedagogy? *Design and Technology Education: An International Journal*, 12(1), 7-9.
- [9] Company, P., Contero, M., & Salvador-Herranz. (July, 2013). Testing rubrics for assessment of quality in CAD modelling. *Proceedings of the Research in Engineering Education Symposium, Kuala Lumpur, Malaysia, July 4-6, 2013*.
- [10] Delahunty, T., Seery, N., & Lynch, R. (2012). An evaluation of the assessment of graphical education at junior cycle in the Irish system. *Design and Technology Education: An International Journal*, 17(2), 9-20.
- [11] Devine, K. D., & Laingen, M. A. (October, 2013). Assessing design intent in an introductory-level engineering graphics course. *Paper published in the proceedings of the 68th Midyear Conference of the Engineering Design Graphics Division of the American Society for Engineering Education, Worcester, Massachusetts, October 20-22, 2013*.
- [12] Dobelis, M., Branoff, T., Nulle, I. Assessment of the Engineering Graphic Literacy Skills. In: *Engineering Graphics BALTGRAF 2013: Scientific Proceedings of the 12th International Conference on Engineering Graphics: The 12th International Conference on Engineering Graphics BALTGRAF 2013, Latvia, Rīga, 5-7 June, 2013*. Riga: Riga Technical University, 2013, pp.69-80. ISBN 9789934507304.
- [13] Giesecke, F. E., Mitchell, A., Spencer, H. H., Hill, I. L., Loving, R. O., Dygdon, J. T., & Novak, J. E. (2000). *Engineering Graphics*. 8th Edition. Upper Saddle River, New Jersey: Pearson, Prentice Hall.
- [14] Hartman, N. W. (2003). Towards the definition and development of expertise in the use of constraint-based CAD tools: Examining practicing professionals (Unpublished doctoral dissertation). North Carolina State University, Raleigh, North Carolina.
- [15] Hartman, N. W. (2009). The development of expertise in the use of constraint-based CAD tools: Examining practicing professionals. *Engineering Design Graphics Journal*, 68(2), 14-26.
- [16] Hartman, N. W. (2009). Defining expertise in the use of constraint-based CAD tools by examining practicing professionals. *Engineering Design Graphics Journal*, 69(1), 6-15.
- [17] Peng, X., McGary, P., Johnson, M., Yalvac, B., & Ozturk, E. (2012). Assessing novice CAD model creation and alteration. *Computer-Aided Design & Applications, PACE*, (2), 9-19.
- [18] Rynne, A., Gaughran, W. F., & Seery, N. (2010). Defining the variables that contribute to developing 3D CAD modelling expertise. In E. Norman & N. Seery (Eds.), *Graphicacy and Modelling*. The International Conference on Design and Technology Educational Research and Curriculum Development, Loughborough, U.K. 161-233.
- [19] Steinhauer, H. M. (2012). Correlation between a student's performance on the Mental Cutting Test and their 3D parametric modeling ability. *Engineering Design Graphics Journal*, 76(3), 44-48.
- [20] Wiebe, E. N., Branoff, T. J., & Hartman, N. W. (2003). Teaching geometry through dynamic modeling in introductory engineering graphics. *Engineering Design Graphics Journal*, 67 (2), 12-20.

ABOUT THE AUTHORS

1. Theodore J. Branoff, Ph.D. is professor and chair of the Department of Technology at Illinois State University. He is the immediate Past-President of ISGG. Dr. Branoff's research interests include spatial visualization and the effects of online instruction for preparing technology education teachers and engineers. Along with teaching courses in introductory engineering graphics, computer-aided design, descriptive geometry, and instructional design, he has conducted CAD and geometric dimensioning & tolerancing workshops for both high school teachers and local industry. His e-mail and postal address are as follows: tjbrano@illstu.edu Department Technology, College of Applied Science & Technology, Illinois State University, P.O. Box 5100, Normal, Illinois, USA, 61790-5100.

2. Modris Dobelis. Ph.D., Dr.sc.ing., professor, head of dep. of Computer Aided Engineering Graphics at the Riga Technical University, Riga, Latvia. His research interests are Computer Aided Design in architecture, civil and mechanical engineering. Along with teaching courses in introductory engineering graphics and descriptive geometry he has provided seminars for high school teachers of technical graphics courses. He can be reached by e-mail: Modris.Dobelis@rtu.lv or through postal address: Dep. of Computer Aided Engineering Graphics, Riga Technical University, Āzenes iela 16/20 - 438, LV-1048, Riga, Latvia.

A STUDY ON MORPHOLOGICAL INTERPRETATION OF THE FAÇADE DESIGN AND FORM DESIGN IN ARCHITECTURE

Kazuma WADA, Hirokazu ABE and Kensuke YASUFUKU

Osaka University

ABSTRACT: This paper explores possible architectural designs of buildings in Japan. We interpreted the impression of the façade design and the form design of buildings as an intuitive representative figure. We classified actual buildings by their fragmented features and analyzed the association between the façade design or the form design and the attribute information of a building. As a result, we found that the building appearance has some relevance to building use and that are some façade designs and form designs are often used as the building design.

Keywords: Façade design, Form design, Use

1. INTRODUCTION

1.1 Background

The general perception of the building appearance of architecture has been changing during the past century.

In the 1900s, the modernist movement in architecture occurred in western countries. In this movement, slogans such as “form follows function” were widespread. This was different from previous ideas. For example, Gropius [1] said “Form of building is led by not itself, but the essence of building, that is, a required function. This is the result of spiritual change and deepening, as well as new technique.” Consequently, this idea led to a decline in the importance of façade design.

In the 1980s, however, the modernist movement was criticized for its uniformity and monotonousness. Venturi [2] said “Modernism architecture devotedly divided pure architectural elements on space, structure and plan, therefore its expression was vacant and bored. Consequently, this became irresponsible and dull expressionism.” He suggested that architecture abandoned the meaning and symbolism of its form.

After 2000, building appearance was reconsidered as an element of architecture. It is now thought that the exterior design should convey the impression of a building to the people. With this trend, Monnai [3] said from viewpoint of a semiotic approach that “The element of building design is also regarded as a sign which reminds us of some concepts.”

In this study, we interpreted the impression of façade design and the form design of buildings again as an intuitive representative figure. We classified buildings by their fragmented features and analyzed the association between the façade design or the form design and the attribute information of building. In addition, we explored the current trends of architectural designs of buildings to predict possible future designs.

1.2 Previous studies

We discuss some relevant previous studies. Ishii et al. [4] chose two buildings at Hosei University as subjects of research and then analyzed their elevation surface above ground. They clarified the features composing the elevation surface.

Nakai et al. [5] researched commercial buildings that faced a street. They examined a feature

and its formal style of expression in building appearance, which consists of a façade design and the form of the volume. Then, they researched patterns of the combination of the units of the façade design and the form of the volume, and revealed typical types and characteristics of these through analysis. However, this study dealt only with commercial buildings and did not represent the architectural design of other types of buildings.

Regarding façade design, a previous study used a semiotic approach. Moriyama et al. [6] analyzed relations between factors integrated into a traditional townscape in Kyoto. They identified the characteristics of Kyoto's design code. In their research, they examined a townscape only. They did not deal with one building. In our paper, we examine the recent trends of the building appearance in architecture in Japan. We also analyze the association between basic information and the two elements of building appearance, the façade and the form. Then, we explore a possible future of building designs.

2. METHOD

2.1 Subjects of research

In this paper, we focus on building photographs published in a popular architectural journal in Japan, called "Shinkenchiku". We extracted subjects of research that were published from January, 2009 to May, 2013 to investigate the trend of building appearance in architecture. The main façade design and form of the subject buildings were studied from the journal photographs. We obtained 219 cases as subjects of research and also obtained basic information about the subjects. The basic information included the total floor area, the number of floors above ground, the completion year, the main use and the location of the site.

2.2 Interpretation of building appearance

Building appearance in architecture consists of many complex factors, such as frames, windows, louvers and external stairs. We interpreted each

subject as an intuitive representative figure to easily understand its composition. Regarding interpretation, we considered two points of view as main factors of building appearance. One is Façade design as the superficial design of buildings. And the other is Form design as the morphological design.

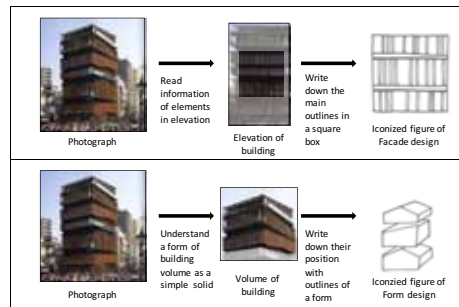


Figure 1: Process of interpretation

As for Façade design, we recognized the principal façade of the building from a photograph in the journal and read information of the elements. We extracted the main outlines of the elements, which consist of Façade design, such as frames, slabs, windows and braces. We wrote the outlines in a square box and made representative models of the Façade design. An example of the interpretation process about Façade design is shown in the upper illustration of Figure 1. Examples of representative figures of the Façade design are shown in Figure 2.

As for Form design, we understood the form of the building volume as a simple solid. We extracted the main outlines of the building form, such as the volume, roof and slab. If a building had multiple volumes, we wrote down their position with outlines of the form. An example of the interpretation process of Form design is shown in the lower illustration of Figure 1. Examples of representative figures of Form design are shown in Figure 3.

2.3 Classification of representative figures

We classified representative figures by their similarity of design. While classifying the Façade design, we paid attention to some elements about the representative figures, such as the direction of outlines, the shape of outlines, the number and size of windows and the pattern of façade design. Then, we camp up with 10 types of Façade design. The types of Façade design are Lattice, Horizontal line, Vertical line, Louver, Window, Pattern, Plane, Curved line, Addition and Gable. Lattice is distinct vertical and horizontal outlines. Horizontal line is distinct horizontal outlines and Vertical line is distinct vertical outlines. Louver is narrow with small outlines of a louver, and Window is outlines of

windows. Pattern is a certain design that is repeatedly used on the façade. Plane is a few outlines on the Façade. Curved line is curved outlines on the Façade. Addition is some outlines of a stair or fence on the façade. Gable has outlines of a delta roof. Representative examples of the types of Façade design are shown in Figure 4.

As for Form design, first we classified the models into Singular type or Plural type according to the number of volumes. Moreover, Plural type was classified into five types by the positions of volumes. The types of positions are Standing in a line, Overlapping, Left and right, Covering and Penetrating. Standing in a line is building volumes situated sideways. Overlappi-

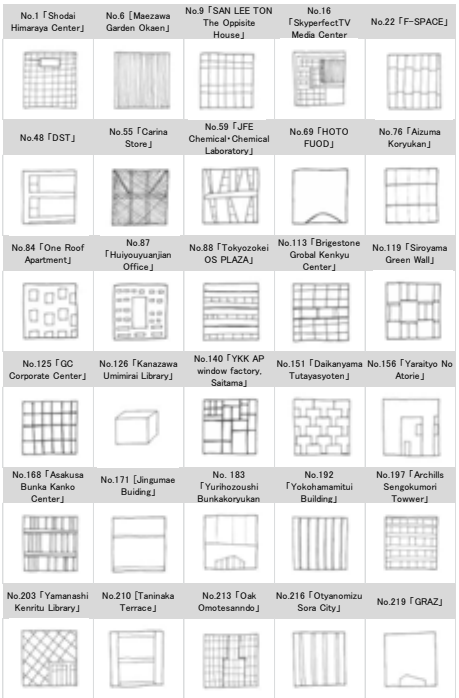


Figure 2: Representative figures of Façade design

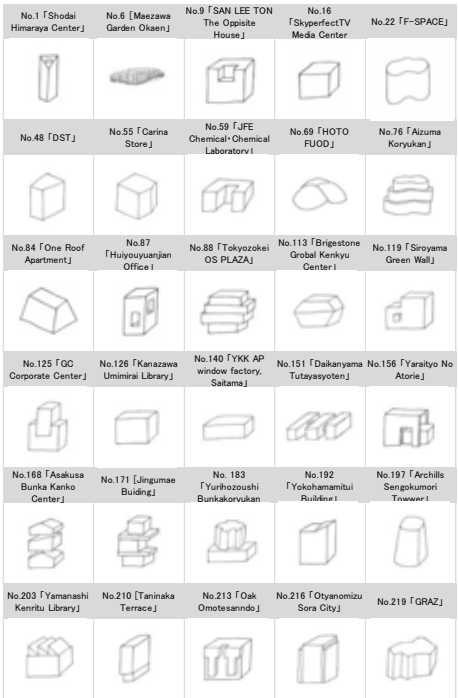


Figure 3: Representative figures of Form design











Lattice	Horizontal line	Vertical line
		
173 UOHDA Hiroyuki THE PLACE for Change Working	67 Y Building	216 Ohtanomizu Sora City
Window	Pattern	Louver
		
103 Prak House Kijoji OIKOS	151 Daikanyama Tutsayasoten	215 Lazona Kawasaki Toshiba Building
Plane	Curved line	Addition
		
219 GRAZ	71 Extension for the fair of Barcelona Gran Via Venue	109 HHH
Gable		
		
212 GINZA KABUKIZA		

Figure 4: Examples of Façade design types








Singular type	Plural type	
		
162 AKASAKA K-TOWER	26 Isunoki	
Standing in a line	Overlapping	Left and right
		
25 Ao	76 Aizuma Kouryukan	50 Beijing Bumps
Covering	Penetrating	
		
177 Marunouchi Eiraku Building	145 Kinki University Department of agriculture laboratory 2	

Figure 5: Examples of position types

ng is volumes piled vertically. Left and right are volumes situated vertically and horizontally. Covering is a large volume that covers a small one. Penetrating is one volume sticking into another volume. Representative examples of the types of positions are shown in Figure 5.

Moreover, we classified models including the Singular type and the Plural type into nine types










Square	Concave	Convex
		
10 Metashutter Office	114 Tyosu Cubic Garden	149 Minatokuribu Takawakodomotyukousei Plaza
Polygon	Curved surface	Delta roof
		
187 Apartment of Numabukuro	107 Apartment of Higashihutyu	17 SOLA Ueno Gas Kaneyama Office
Flat roof and slab	Wall	Column and beam
		
186 Komatunagi Terrace	32 OYM	6 Maszawa Garden Okaen

Figure 6: Examples of Form design types

by the form of the volume. The types of forms are Square, Concave, Convex, Polygon, Curved surface, Delta roof, Flat roof and slab, Wall, Column and beam. Square is a rectangular parallelepiped volume. In Concave, a part of the volume is sunk in. In Convex, a part of the volume sticks out. Polygon is the shape of a polyhedron. Curved surface is a part of the volume consisting of a curved surface. Delta roof is a pitched roof. Flat roof and slab is a horizontal flat volume. Wall is a vertical plane volume. Column and beam is volumes of a bar just like a column or a beam. Representative examples of the types of forms are shown in Figure 6.

3. RESULTS

We analyzed each type mentioned in the previous section and investigated the relation between each type and the basic information of building. For each type of Façade design, the rate of the number of new buildings built from 2008 to 2012 is shown in Figure 7. As shown in Figure 7, there are a lot of buildings with Façade design of Lattice or Window. About each type of Form design, the rate of the number of new buildings which was built from 2008 to

2012 is shown in Figure 8. As shown in Figure 8, the rate of Square has been increasing in recent years.

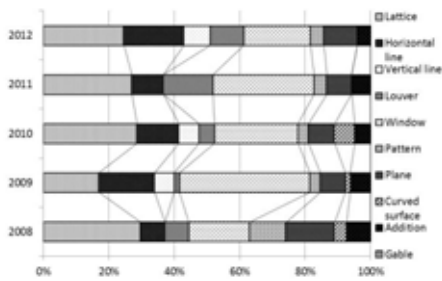
Table 1 shows the rate of the number of buildings with each Façade design to each type of Form design. As shown in Table 1, the score in the Façade design type of Lattice and Window were higher than that of the other types. This suggests that the type of Lattice and Window types have versatility. The ratio of the type of Horizontal line was high in the Form design type of Roof and slab, and that of Plane was high in the Form design type of Wall.

Next, we used Hyashi's quantification theory type III [7] to analyze the degree of similarity

between the Façade design and the Form design, each type and the main use of the building. In this method, similar items are closely plotted in the graph.1

Figure 9 shows the graph about the similarity between the Façade design and the Form design obtained as an analysis result. As shown in Figure 9, Plane is relevant to Wall and Curved line is relevant to Curved surface.

Figure 10 shows the graph of the similarity between the main use of a building and the Façade design obtained from the result of the analysis. As shown in Figure 10, the use as an office was strongly relevant to the Façade design of Lattice and the use as housing was related to the Façade design of Window and Addition. In addition-



Façade design type 2008-2012

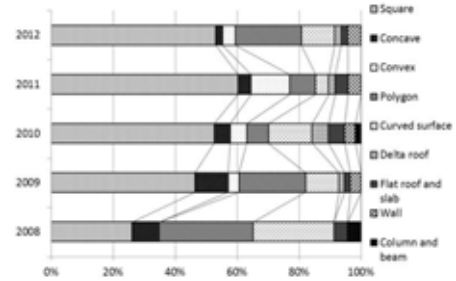


Figure 8: The rate of new buildings for each Form design type 2008-2012

Table 1: The rate of buildings with each Façade design type about Form design type

	Percentage(%)	Lattice	Horizontal line	Vertical line	Louver	Window	Pattern	Plane	Curved line	Addition	Gable
Square	100	25.3	13.0	5.8	8.4	24.0	5.2	9.7	2.6	5.2	0.6
Concave	100	33.3	20.0	0.0	0.0	40.0	0.0	0.0	0.0	6.7	0.0
Convex	100	37.5	6.3	6.3	12.5	37.5	0.0	0.0	0.0	0.0	0.0
Polygon	100	26.7	13.3	6.7	2.2	35.6	2.2	11.1	0.0	2.2	0.0
Curved surface	100	13.9	11.1	2.8	5.6	22.2	11.1	11.1	11.1	11.1	0.0
Delta roof	100	25.0	0.0	12.5	0.0	50.0	0.0	0.0	0.0	0.0	12.5
Flat roof and slab	100	0.0	20.0	0.0	20.0	40.0	0.0	0.0	0.0	20.0	0.0
Wall	100	0.0	7.1	0.0	0.0	21.4	0.0	50.0	0.0	21.4	0.0
Column and beam	100	50.0	0.0	0.0	50.0	0.0	0.0	0.0	0.0	0.0	0.0

on, the use of commerce and public facilities has relevance to the Façade design of the Horizontal line, Vertical line, Curved line, and Plane.

Figure 11 shows the graph for the similarity between the main use of a building and the Form design obtained as an analysis result. As shown in Figure 11, the use of an office and commerce were relevant to the Form design of Square and the use of housing was similar to the Form design of Polygon and Convex. The use of public facilities was related to the Form design of Wall and Curved surface.

Consequently, we obtained some repeated combinations of the building appearance and the main use of a building. Such combinations are often seen in Japan. Figure 12 shows the representative examples with the Façade design the

Form design and the main use of the building. Their factors have a strong relevance to each other.

We identified that the number of buildings having the Form design of Square and the Façade design of Lattice or Window have been increasing in recent years. Regarding the relations between building appearance and main use, it was found that a building of office and commerce use often has the Façade design of Lattice, Vertical line or Horizontal line. The use of housing frequently has the Façade design of Window. Moreover, the use of public facilities has various Façade designs and Form designs. This suggests that these buildings were designed by using various design elements in order to attract people.

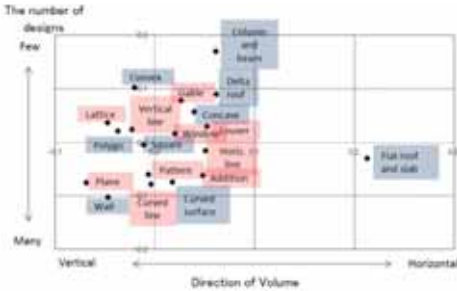


Figure 9: Similarity between Façade design and Form design

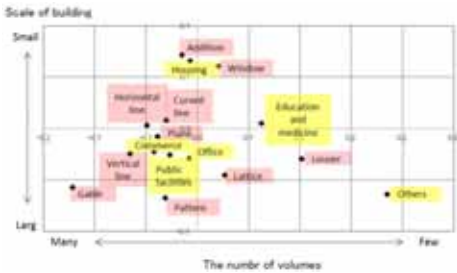


Figure 10: Similarity between Façade design and use

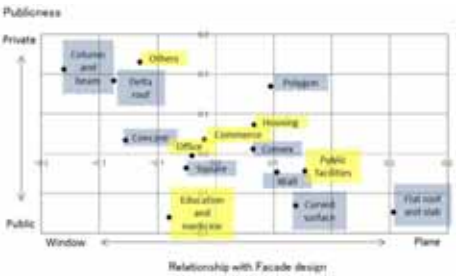


Figure 11: Similarity between Form design and use

Name	122 Toyosu Front	110 Forum Building
Photograph		
Main use	Office	Office
Position of volumes	Singular type	Singular type
Façade design type	Lattice	Lattice
Form design type	Square	Square
187 Apartment of Numabukuro	134 kap	115 UEHONMACHI YUFURA
Housing	Housing	Commerce
Singular type	Singular type	Standing in a line
Window	Window	Horizontal line + Vertical line
Polygon	Convex	Square

Figure 12: Examples of typical combinations building appearance and use

4. CONCLUSION

In this paper, the results are as follows.

1. The number of buildings with the Façade design of a Lattice and Window and the Form design of a Square has been increasing in recent years. These types of Façade design and Form design are often seen all over Japan.
2. Regarding the relation between Façade design and Form design, the ratio of the Façade design of the Lattice and Window was higher than that of other types of Façade design in comparison with all Form design types. In brief, Lattice and Window tend to be easily used as a façade design in spite of the type of building form.
3. As a result of the analyses in this paper, we found that some combinations with Façade design and Form design are frequently used for buildings with certain uses. Table 5 shows a list of the relevance between main use and the elements of building appearance.

4. For the purpose of developing diversity of the building appearance in the future, we need to pay attention to combinations that have a weak relation. As a result, the conventional relation between building appearance and the use of building would cease.

REFERENCES

- [1] W. Gropius, "Internationale Architektur," Gebrueder Mann Verlag, 1998
- [2] R. Venturi, "Leaning of Las Vegas," The MIT Press, 1977.
- [3] T. MONNAI, "Semiotic Approach to Architecture, City and Development of Townscape Semiotic," Design Research a special number, pp. 13-29, 2002.
- [4] S. ISHII, N. ANDO and M. TANEDA, "A Study on the Façade of Modernism Architecture: Case Study of Hosei Univ 55/58nenkan," Journal for Geometry and Graphics, Summaries of Technical Papers of Annual Meeting, Graphic Science of Japan, pp. 137-142, 2011.

Table 2 Relevance

Main use	Scale of building	The number of volumes	The position of volumes	Façade design		Form design	
Office	Large	Plural	Overlapping	Lattice Vertical line	Horizontal line	Square Concave Convex	
Housing	Small	Singular		Window	Addition	Polygon Square	Convex
Commerce	Small-Large		Standing in a line	Vertical line Addition	Horizontal line Plane	Square Convex	Concave
Public facilities	Medium-large		Overlapping	Vertical line Curved line	Horizontal line Plane	Curved surface	Wall
Education and Medicine	Small			Window	Louver	Square	
Others	Small-Medium		Penetrating	Pattern	Plane	Column and beam Square	

- [5] K. NAKAI, R. NEMOTO and K. SAKAMOTO, "Exterior Spatial Composition of Contemporary Japanese Commercial Buildings: A study on architectural composition with articulation of exterior volumes (5)," *Journal of Architecture and Planning*, pp. 65-70, 2007.
- [6] M. Moriyama and T. Monnai, "Description of Systemic Code and Sign Network of Townscapes in Kyoto: Analysis of the Relation Design in Townscapes (part1)," *Journal of Architecture and Planning*, pp.1506-1507, 2010.
- [7] K. Toshiharu, "Relatedness Evaluation Qualitative Analysis," *Journal of Society of Humanities Sapporo Gakuin University*, No.83, pp. 61-100, 2008.

ABOUT THE AUTHORS

1. Kazuma WADA, BA. (Eng.) is a student
Osaka University, 2-1 Ymadaoka, Suita,
Osaka 565-0871, Japan
e-mail: wada_kazuma@arch.eng.osaka-u.ac.jp
2. Hirokazu ABE, Dr. (Eng.), is a professor at
Osaka University, Cyber Media Center,
5-1Mihogaoka, Ibaraki, Osaka 567-0047,
Japan
e-mail: abe@arch.eng.osaka-u.ac.jp
3. Kensuke YASUFUKU, Dr. (Eng.), is an as-
sistant professor at Osaka University, Cyber
Media Center, 5-1Mihogaoka, Ibaraki, Osaka
567-0047, Japan
e-mail: yasufuku@arch.eng.osaka-u.ac.jp

A STUDY ON SPATIAL CYCLOID GEARING

Giorgio FIGLIOLINI¹, Hellmuth STACHEL², and Jorge ANGELES³

¹University of Cassino & Southern Lazio, Italy ²Vienna University of Technology, Austria

³McGill University, Montréal, Canada

ABSTRACT: Understanding the geometry of gears with skew axes is a complex task, hard to grasp and to visualize. However, due to Study's Principle of Transference, the geometric treatment based on dual vectors can be readily derived from that of the spherical case. This paper is based on Martin Disteli's work and on the authors' previous results where Camus' concept of an auxiliary curve is extended to the case of skew gears. We focus on the spatial analogue of the following case of cycloid bevel gears: When the auxiliary curve is specified as a pole tangent, we obtain 'pathologic' spherical involute gears with vanishing pressure angle. The profiles are always penetrating at the meshing point because of G^2 -contact.

In view of the Camus Theorem, the spatial analogue of the pole tangent is a skew orthogonal helicoid Π_4 as auxiliary surface. Its axis lies on the cylindroid and is normal to the instant screw axis (ISA). Under the roll-sliding of Π_4 along the axodes Π_2 and Π_3 of the gears, any generator g of Π_4 traces a pair of conjugate flanks Φ_2, Φ_3 with permanent line contact. Again, these flanks are not realizable because of the reasons below:

(1) When g coincides with the ISA, the singular lines of the two flanks come together. At each point of g the two flanks share the tangent plane, but in the case of external gears the surfaces open toward opposite sides.

(2) We face the spatial analogue of a spherical G^2 -contact, which surprisingly does not mean a G^2 -contact at all points of g but only at a single point combined with a mutual penetration of the flanks Φ_2 and Φ_3 .

However, when instead of a line g a plane Φ_4 is attached to the right helicoid Π_4 , the envelopes of Φ_4 under the roll-sliding of Π_4 along Π_2 and Π_3 are torses that serve as conjugate tooth flanks Φ_2, Φ_3 with a permanent line contact. So far, it seems that these flanks, Φ_2 and Φ_3 , are geometrically feasible. This is a possible spatial generalization of octoidal gears or even of planar involute gears.

Keywords: Gears with skew axes, Cycloidal gearing, Involute gearing, Cylindroid, Camus Theorem

1. INTRODUCTION

Let the motions of two gears Σ_2, Σ_3 , against the gear box Σ_1 be given, i.e., the rotations $\Sigma_2/\Sigma_1, \Sigma_3/\Sigma_1$ about fixed skew axes p_{21} and p_{31} with angular velocities ω_{21}, ω_{31} , respectively. The dual unit vectors representing the axes p_{21} and p_{31} are denoted by $\hat{\mathbf{p}}_{21}$ and $\hat{\mathbf{p}}_{31}$, respectively. We use a Cartesian coordinate frame $\mathcal{F}(O; x_1, x_2, x_3)$ with $\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2$ denoting the dual unit vectors of the x_1 - and x_2 -axis. The given axes p_{21} and p_{31} of the wheels are assumed to be symmetrically

placed with respect to the x_1 -axis such that the x_3 -axis is the common normal of the gear axes.

Using the *dual angle* $\hat{\alpha} = \alpha + \varepsilon\alpha_0$ between the x_1 -axis and p_{21} , we can set (see Fig. 1)

$$\begin{aligned}\hat{\mathbf{p}}_{21} &= \cos \hat{\alpha} \hat{\mathbf{e}}_1 - \sin \hat{\alpha} \hat{\mathbf{e}}_2, \\ \hat{\mathbf{p}}_{31} &= \cos \hat{\alpha} \hat{\mathbf{e}}_1 + \sin \hat{\alpha} \hat{\mathbf{e}}_2.\end{aligned}\quad (1)$$

We limit ourselves to the skew case and assume

$$0 < \alpha < \pi/2 \text{ and } \alpha_0 \neq 0 \quad (2)$$

though most of the arguments hold also in the spherical case, $\alpha_0 = 0$, and in the planar case, $\alpha = 0$, with parallel axes.

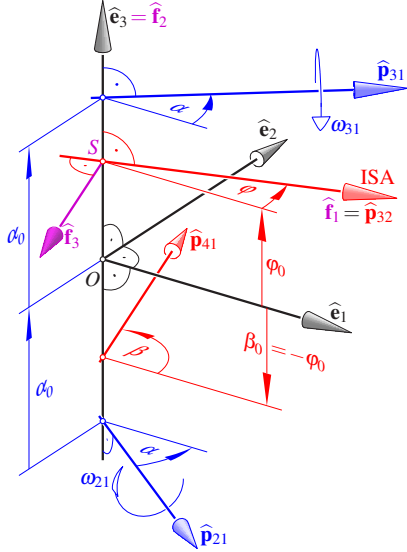


Figure 1: Skew axes \hat{p}_{21} , \hat{p}_{31} of the gear wheels, the ISA \hat{p}_{32} and the axis \hat{p}_{41} of the auxiliary surface $\Pi_4 \subset \Sigma_4$ in the particular case $\beta = \varphi + \pi/2$. The oriented lines $\hat{f}_1, \hat{f}_2, \hat{f}_3$ form the Frenet frame of the axodes. This frame remains fixed in the gear frame Σ_1 .

In addition, let $\hat{\varphi}$ denote the dual angle between \hat{e}_1 and the ISA, i.e., the relative axis \hat{p}_{32} . Then we obtain

$$\begin{aligned} \hat{p}_{32} &= \cos \hat{\varphi} \hat{e}_1 + \sin \hat{\varphi} \hat{e}_2 \text{ and} \\ \hat{\omega}_{32} \hat{p}_{32} &= \omega_{31} \hat{p}_{31} - \omega_{21} \hat{p}_{21}. \end{aligned} \quad (3)$$

The comparison of coefficients and [4, Eq. (7)] lead to

$$\begin{aligned} \tan \varphi &= \frac{\omega_{31} + \omega_{21}}{\omega_{31} - \omega_{21}} \tan \alpha \text{ and} \\ \varphi_0 &= R \sin 2\varphi \text{ with } R = \frac{\alpha_0}{\sin 2\alpha}. \end{aligned} \quad (4)$$

The vector product of both sides of the last equation in (3) with \hat{p}_{21} and \hat{p}_{31} (compare [6,

Eq. (12)]) results in

$$\frac{\omega_{21}}{\sin(\hat{\varphi} - \hat{\alpha})} = \frac{\omega_{31}}{\sin(\hat{\varphi} + \hat{\alpha})} = \frac{\hat{\omega}_{32}}{\sin 2\hat{\alpha}}, \quad (5)$$

which sometimes is called the *dual Sine-Theorem* applied to the dual ‘triangle’ $\omega_{21} \hat{p}_{21}$ and $\omega_{31} \hat{p}_{31}$ and $\hat{\omega}_{32} \hat{p}_{32}$. This implies

$$\hat{\omega}_{32} = \frac{\omega_{21} \sin 2\hat{\alpha}}{\sin(\hat{\varphi} - \hat{\alpha})} \quad (6)$$

and, consequently, for the pitch of the relative motion Σ_3/Σ_2 like [4, Eq. (15)]

$$\begin{aligned} h_{32} &= \frac{\omega_{320}}{\omega_{32}} = R(\cos 2\alpha - \cos 2\varphi) \\ &= 2R(\cos^2 \alpha - \cos^2 \varphi). \end{aligned} \quad (7)$$

The axodes of the relative motion Σ_3/Σ_2 are one-sheet hyperboloids $\Pi_3 \subset \Sigma_3$ and $\Pi_2 \subset \Sigma_2$, swept by the relative axis p_{32} under the inverse rotations Σ_1/Σ_2 and Σ_1/Σ_3 about p_{21} and p_{31} , respectively.

2. THE SPATIAL CAMUS THEOREM

The following lemma was first published by Disteli (see [6, Theorem 2] and the references therein).

Lemma 1. *For given wheels Σ_2, Σ_3 there exists a frame Σ_4 such that the screws of $\Sigma_4/\Sigma_2, \Sigma_4/\Sigma_3$ and Σ_3/Σ_2 are equal at every instant if and only if the instant axis p_{41} of Σ_4/Σ_1 is located on the Plücker conoid Ψ , but different from p_{32} .*

Let $\hat{\beta}$ be the dual angle between between the x_1 -axis and \hat{p}_{41} (Fig. 1). Then we can write

$$\hat{p}_{41} = \cos \hat{\beta} \hat{e}_1 + \sin \hat{\beta} \hat{e}_2. \quad (8)$$

If we specify $p_{41} \subset \Psi$ different from p_{21}, p_{31}, p_{32} , then $\varphi \neq \pm\alpha, \beta$. From the equation [6, Eq. (8)], which defines the Plücker conoid, we obtain

$$\beta_0 = R \sin 2\beta. \quad (9)$$

The dual Sine-Theorem applied to the triangle $\omega_{21} \hat{p}_{21}$, $\hat{\omega}_{41} \hat{p}_{41}$ and $\hat{\omega}_{42} \hat{p}_{42} = \hat{\omega}_{42} \hat{p}_{32}$ gives (compare [6, Eq. (12)])

$$\frac{\omega_{21}}{\sin(\hat{\varphi} - \hat{\beta})} = \frac{\hat{\omega}_{41}}{\sin(\hat{\varphi} + \hat{\alpha})} = \frac{\hat{\omega}_{42}}{\sin(\hat{\alpha} + \hat{\beta})} \quad (10)$$

and, analogously, for $\omega_{31}\hat{\mathbf{p}}_{31}$, $\hat{\omega}_{41}\hat{\mathbf{p}}_{41}$ and $\hat{\omega}_{43}\hat{\mathbf{p}}_{43} = \hat{\omega}_{43}\hat{\mathbf{p}}_{32}$

$$\frac{\omega_{31}}{\sin(\hat{\varphi} - \hat{\beta})} = \frac{\hat{\omega}_{41}}{\sin(\hat{\varphi} - \hat{\alpha})} = \frac{\hat{\omega}_{43}}{\sin(\hat{\beta} - \hat{\alpha})}. \quad (11)$$

The instant pitch $h_{41} = \omega_{410}/\omega_{41}$ is defined by [6, Eq. (9)] as

$$h_{41} = \frac{\omega_{410}}{\omega_{41}} = R(\cos 2\alpha - \cos 2\beta). \quad (12)$$

Let Π_4 be the ruled helical surface¹ traced by the relative axis p_{32} under the helical motion Σ_1/Σ_4 about p_{41} with pitch h_{41} . We call Π_4 the *auxiliary surface* (for further details see [5]). It forms together with Π_2 and Π_3 the axodes of the relative motions of Σ_4 against Σ_2 and Σ_3 , i.e., the motions Σ_4/Σ_2 and Σ_4/Σ_3 are defined by the rolling and sliding of Π_4 along the hyperboloids Π_2 and Π_3 , respectively.

The importance of the auxiliary surface $\Pi_4 \subset \Sigma_4$ lies in [6, Theorem 3] which we recall as below:

Theorem 2. [Spatial Camus Theorem]

For any line g attached to Σ_4 , the surfaces Φ_2 , Φ_3 traced by g under the relative motions Σ_4/Σ_2 and Σ_4/Σ_3 , respectively, are conjugate tooth flanks of Σ_3/Σ_2 . At any instant, the meshing points for these flanks are located on a straight line.

With respect to the gear frame Σ_1 , the locus of the meshing lines, i.e., the *meshing surface* or *surface of action*, is traced by g under Σ_4/Σ_1 with the fixed twist $\hat{\mathbf{q}}_{41} = \hat{\omega}_{41}\hat{\mathbf{p}}_{41}$. Consequently, it is a helical surface with axis $\hat{\mathbf{p}}_{41}$.

3. THE DISTELI AXES OF A RULED SURFACE

Along each non-torsal generator g of a ruled surface a *Frenet frame* can be defined, consisting of: g itself; the *central normal* n , which is the surface normal at the striction point; and the *central*

tangent t (see, e.g., [1, 2]). This triplet of mutually orthogonal axes meets at the *striction point* S of g , defined on the *striction curve* (Fig. 2). The central tangent is orthogonal to the asymptotic plane and tangent to the surface at the striction point.

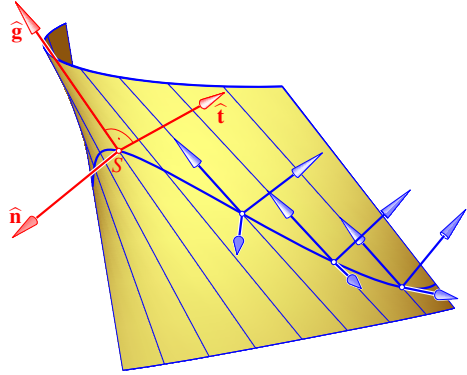


Figure 2: Frenet frame $(\hat{\mathbf{g}}, \hat{\mathbf{n}}, \hat{\mathbf{t}})$ and striction curve of a ruled surface.

Let, in dual-vector notation², the ruled surface be given by the twice-differentiable dual vector function $\hat{\mathbf{g}}(t)$, $t \in I$. Then, the derivatives of the Frenet frame $(\hat{\mathbf{g}}, \hat{\mathbf{n}}, \hat{\mathbf{t}})$ satisfy the *Frenet equations*—Eq. (10) of [1]—namely,

$$\begin{aligned} \hat{\mathbf{g}}' &= \hat{\lambda} \hat{\mathbf{n}} &= \hat{\mathbf{q}} \times \hat{\mathbf{g}} \\ \hat{\mathbf{n}}' &= -\hat{\lambda} \hat{\mathbf{g}} &+ \hat{\mu} \hat{\mathbf{t}} &= \hat{\mathbf{q}} \times \hat{\mathbf{n}} \\ \hat{\mathbf{t}}' &= -\hat{\mu} \hat{\mathbf{n}} &= \hat{\mathbf{q}} \times \hat{\mathbf{t}} \end{aligned} \quad (13)$$

with $\hat{\mathbf{q}} = \hat{\mu} \hat{\mathbf{g}} + \hat{\lambda} \hat{\mathbf{t}} = \hat{\omega} \hat{\mathbf{g}}^*$,

$\hat{\mathbf{g}}^*$ with $\hat{\mathbf{g}}^* \cdot \hat{\mathbf{g}}^* = 1$ being the *Disteli axis* and $\hat{\omega}^2 = \hat{\lambda}^2 + \hat{\mu}^2$, provided $\hat{\lambda} \neq 0$. By the last condition we exclude stationary (= singular) generators.

The Frenet equations (13) contain two dual coefficients, $\hat{\lambda} = \lambda + \varepsilon \lambda_0$ and $\hat{\mu} = \mu + \varepsilon \mu_0$. Various formulas expressing invariants of the ruled

¹ In this paper the term ‘ruled surface’ stands for a twice continuously differentiable one-parameter set of oriented lines.

²From now on we identify oriented lines with their dual unit vector—with respect to any well-defined coordinate frame. In this sense we speak of the ‘line $\hat{\mathbf{g}}$ ’.

surface in terms of λ , λ_0 , μ , and μ_0 can be found in [1, Theorems 1–3].³ Here we adopt a different approach.

The dual representation $\widehat{\mathbf{g}}(t) = \mathbf{g}(t) + \varepsilon \mathbf{g}_0(t)$, $t \in I$, of the ruled surface gives rise to a real parametrization, namely

$$\mathbf{x}(t, u) = [\mathbf{g}(t) \times \mathbf{g}_0(t)] + u \mathbf{g}(t), \quad (t, u) \in I \times \mathbb{R}. \quad (14)$$

Here we recall that $\mathbf{g} \times \mathbf{g}_0$ is the position vector of the pedal point of the generator $\widehat{\mathbf{g}}$ with respect to the origin of the underlying coordinate frame. The derivatives

$$\begin{aligned} \frac{d}{dt} \widehat{\mathbf{g}} &= \widehat{\dot{\mathbf{g}}} = \dot{\mathbf{g}} + \varepsilon \dot{\mathbf{g}}_0 = \widehat{\lambda} \widehat{\mathbf{n}} \\ &= \lambda \mathbf{n} + \varepsilon (\lambda_0 \mathbf{n} + \lambda \mathbf{n}_0), \\ \frac{d^2}{dt^2} \widehat{\mathbf{g}} &= \widehat{\ddot{\mathbf{g}}} = \ddot{\mathbf{g}} + \varepsilon \ddot{\mathbf{g}}_0 = -\widehat{\lambda}^2 \widehat{\mathbf{g}} + \widehat{\dot{\lambda}} \widehat{\mathbf{n}} + \widehat{\lambda} \widehat{\mu} \widehat{\mathbf{t}} \\ &= -\lambda^2 \mathbf{g} + \dot{\lambda} \mathbf{n} + \lambda \mu \mathbf{t} + \varepsilon (-2\lambda \lambda_0 \mathbf{g} - \lambda^2 \mathbf{g}_0 \\ &\quad + \dot{\lambda}_0 \mathbf{n} + \dot{\lambda} \mathbf{n}_0 + \lambda_0 \mu \mathbf{t} + \lambda \mu_0 \mathbf{t} + \lambda \mu \mathbf{t}_0) \end{aligned} \quad (15)$$

determine the partial derivatives of the parametrization $\mathbf{x}(t, u)$:

$$\mathbf{x}_t = (\dot{\mathbf{g}} \times \mathbf{g}_0) + (\mathbf{g} \times \dot{\mathbf{g}}_0) + u \dot{\mathbf{g}}, \quad \mathbf{x}_u = \mathbf{g}$$

and

$$\begin{aligned} \mathbf{x}_{tt} &= (\ddot{\mathbf{g}} \times \mathbf{g}_0) + 2(\dot{\mathbf{g}} \times \dot{\mathbf{g}}_0) + (\mathbf{g} \times \ddot{\mathbf{g}}_0) + u \ddot{\mathbf{g}}, \\ \mathbf{x}_{tu} &= \dot{\mathbf{g}} = \lambda \mathbf{n}, \quad \mathbf{x}_{uu} = \mathbf{0}. \end{aligned}$$

We study the derivatives at the points of a single generator, say, at $t = 0$. For this purpose we use the triplet $(\widehat{\mathbf{g}}(0), \widehat{\mathbf{n}}(0), \widehat{\mathbf{t}}(0))$ as the new coordinate frame; now the striction point $\mathbf{s}(0)$ of $\widehat{\mathbf{g}}(0)$ is the origin of the frame in question. Thus we may set

$$\begin{aligned} \mathbf{g}(0) &= \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{n}(0) = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{t}(0) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \\ \mathbf{g}_0(0) &= \mathbf{n}_0(0) = \mathbf{t}_0(0) = \mathbf{0}. \end{aligned}$$

This yields

$$\widehat{\mathbf{g}}(0) = \begin{pmatrix} 0 \\ \lambda \\ 0 \end{pmatrix} + \varepsilon \begin{pmatrix} 0 \\ \lambda_0 \\ 0 \end{pmatrix},$$

³For example: The dual part \mathbf{q}_0 of the twist $\widehat{\mathbf{q}}$ equals the instant velocity vector of the origin \mathbf{s} . Consequently, for the striction σ (see Fig. 3) we get $\tan \sigma = \lambda/\mu$.

$$\widehat{\mathbf{g}}(0) = \begin{pmatrix} -\lambda^2 \\ \dot{\lambda} \\ \lambda \mu \end{pmatrix} + \varepsilon \begin{pmatrix} -2\lambda \lambda_0 \\ \dot{\lambda}_0 \\ \lambda_0 \mu + \lambda \mu_0 \end{pmatrix}$$

and therefore

$$\mathbf{x}_t(0, u) = \begin{pmatrix} 0 \\ \lambda u \\ \lambda_0 \end{pmatrix}, \quad \mathbf{x}_u(0, u) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad (16)$$

$$\begin{aligned} \mathbf{x}_{tt}(0, u) &= \begin{pmatrix} -\lambda^2 u \\ -\lambda_0 \mu - \lambda \mu_0 + \dot{\lambda} u \\ \dot{\lambda}_0 + \lambda \mu u \end{pmatrix}, \\ \mathbf{x}_{tu}(0, u) &= \begin{pmatrix} 0 \\ \lambda \\ 0 \end{pmatrix}, \quad \mathbf{x}_{uu}(0, u) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \end{aligned} \quad (17)$$

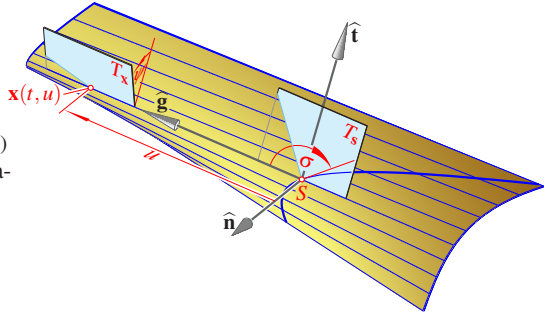


Figure 3: The distribution parameter δ defines the tangent planes T_x along the generator $\widehat{\mathbf{g}}$ by $\tan \psi = -u/\delta$. The angle σ between $\widehat{\mathbf{g}}$ and the striction curve is called the *striction angle* or the *striction*.

The vector product $\mathbf{b} = \mathbf{x}_t \times \mathbf{x}_u$ is a normal vector of the ruled surface, provided the surface point is regular, which means $\mathbf{b} \neq \mathbf{0}$. The coordinates

$$\mathbf{b}(0, u) = \begin{pmatrix} 0 \\ \lambda_0 \\ -\lambda u \end{pmatrix} \quad (18)$$

reveal that at generators with $\lambda \lambda_0 \neq 0$ the angle ψ between the central normal vector $\mathbf{b}(0, 0) = \lambda \mathbf{n}$ and the normal vector $\mathbf{b}(0, u)$ (see Fig. 3) satisfies the equation

$$\tan \psi = \frac{-\lambda u}{\lambda_0} = -\frac{u}{\delta} \quad \text{with } \delta = \frac{\lambda_0}{\lambda}. \quad (19)$$

The quotient δ is called the *distribution parameter*. This is a geometric invariant, i.e., invariant against parameter transformations. Generators with $\lambda_0 = 0$ and hence $\delta = 0$ are called *torsal*: Here all points with $u \neq 0$ have the same tangent plane; the striction point ($u = 0$) is singular because of $\mathbf{b}(0, 0) = \mathbf{0}$.

Cylindrical generators are defined by $\dot{\mathbf{g}} = \mathbf{0}$ or $\lambda = 0$. Here, all points are possible striction points, for which we set $\delta := \infty$.

4. TWO RULED SURFACES WITH LINE CONTACT

For our study on cycloid gearing we need some results concerning the Disteli axes $\hat{\mathbf{g}}^*$ of a ruled surface. According to (13), $\hat{\mathbf{q}} = \hat{\omega} \hat{\mathbf{g}}^*$ is the twist and therefore $\hat{\mathbf{g}}^*$ the instant screw axis of the moving Frenet frame. From Eqs. (15) and (13) follows the relation below:

$$\hat{\mathbf{g}} \times \hat{\mathbf{g}}^* = \hat{\lambda} \hat{\mathbf{n}} \times (-\hat{\lambda}^2 \hat{\mathbf{g}} + \hat{\lambda} \hat{\mathbf{n}} + \hat{\lambda} \hat{\mu} \hat{\mathbf{t}}) = \hat{\lambda}^2 \hat{\omega} \hat{\mathbf{g}}^*. \quad (20)$$

Due to [1, Theorem 3, 3], the dual angle $\hat{\gamma} = \gamma + \varepsilon \gamma_0$ between the generator $\hat{\mathbf{g}}$ and the corresponding Disteli axis $\hat{\mathbf{g}}^*$ satisfies

$$\cot \hat{\gamma} = \frac{\hat{\mu}}{\hat{\lambda}}, \text{ hence} \quad \cot \gamma = \frac{\mu}{\lambda} \text{ and } \gamma_0 = \frac{\lambda \mu_0 - \lambda_0 \mu}{\lambda^2 + \mu^2}. \quad (21)$$

This is a consequence of the two standard products

$$\hat{\mathbf{g}} \cdot \hat{\mathbf{g}}^* = \cos \hat{\gamma} = \frac{\hat{\mu}}{\hat{\omega}}, \quad \hat{\mathbf{g}} \times \hat{\mathbf{g}}^* = \sin \hat{\gamma} \hat{\mathbf{n}} = -\frac{\hat{\lambda}}{\hat{\omega}} \hat{\mathbf{n}},$$

and of the rule that the dual extension of an analytic real function $f(t)$ is defined as $f(\hat{t}) = f(t + \varepsilon t_0) = f(t) + \varepsilon t_0 f'(t)$, which yields

$$\cot \hat{\gamma} = \cot \gamma + \varepsilon \gamma_0 (1 + \cot^2 \gamma).$$

The dual angle between the moving $\hat{\mathbf{g}}(t)$ and the fixed $\hat{\mathbf{g}}^*(0)$ is stationary of order 2 at $t = 0$ (see [1, Theorem 3, 4]). Due to the spheric analogy, $\cot \hat{\gamma}$ can be called the *dual (geodesic) curvature* of the ruled surface.

Lemma 3. *If two ruled surfaces are in contact at all points of a common generator and if they share the corresponding Frenet frame and the Disteli axis, then their dual coefficients in the Frenet equations differ at the corresponding parameter values only by a real factor $c \neq 0$.*

The proof is straightforward and left for the reader.

Theorem 4. *Let $\hat{\mathbf{g}}(t)$ and $\tilde{\hat{\mathbf{g}}}(\tilde{t})$ be two twice-differentiable ruled surfaces which at $t = \tilde{t} = 0$ share the Frenet frame, the distribution parameter $\delta(0) = \tilde{\delta}(0)$ and the Disteli axis. Then, the surfaces have a G^2 -contact at the striction point of the common generator.*

If by Lemma 3 $\hat{\lambda}(0) = c \hat{\lambda}(0)$ and $\tilde{\hat{\mu}}(0) = c \hat{\mu}(0)$, then there is a G^2 -contact at all points of $\hat{\mathbf{g}}(0) = \tilde{\hat{\mathbf{g}}}(0)$ if and only if $\tilde{\delta}(0) = c \delta(0)$.

Proof: The dual vector function $\hat{\mathbf{g}}(t)$ determines the real parametrization $\mathbf{x}(t, u)$ of the ruled surface as presented in (14). The partial derivatives at $t = 0$, as given in (16), define the coefficients of the *first fundamental form* as

$$\begin{aligned} E(0, u) &= \mathbf{x}_t \cdot \mathbf{x}_t = \lambda^2 u^2 + \lambda_0^2, \\ F(0, u) &= \mathbf{x}_t \cdot \mathbf{x}_u = 0, \\ G(0, u) &= \mathbf{x}_u \cdot \mathbf{x}_u = 1. \end{aligned} \quad (22)$$

For the coefficients of the *second fundamental form* we obtain

$$\begin{aligned} L &= \frac{1}{\|\mathbf{b}\|} \mathbf{b} \cdot \mathbf{x}_{tt} = \frac{1}{\sqrt{\lambda_0^2 + \lambda^2 u^2}} \left[-\lambda_0(\lambda_0 \mu + \lambda \mu_0) + (\dot{\lambda} \lambda_0 - \lambda \dot{\lambda}_0)u - \lambda^2 \mu u^2 \right], \\ M &= \frac{1}{\|\mathbf{b}\|} \mathbf{b} \cdot \mathbf{x}_{tu} = \frac{1}{\sqrt{\lambda_0^2 + \lambda^2 u^2}} \lambda \lambda_0, \\ N &= \frac{1}{\|\mathbf{b}\|} \mathbf{b} \cdot \mathbf{x}_{uu} = 0. \end{aligned} \quad (23)$$

For the sake of brevity we skip the detailed analysis, which reveals that at any point $\mathbf{x}(0, u)$ on the common generator $t = 0$ the equations $\tilde{E} = c^2 E$, $\tilde{L} = c^2 L$, and $\tilde{M} = c M$ characterize the G^2 -contact between the two surfaces. \square

5. THE CURVATURE OF THE RULED TOOTH FLANKS

In the realm of gearing, we need two different Frenet frames, the frame $(\hat{\mathbf{f}}_1, \hat{\mathbf{f}}_2, \hat{\mathbf{f}}_3)$ for the axodes with the ISA $\hat{\mathbf{f}}_1$ (see Fig. 1) and the frame $(\hat{\mathbf{g}}_1, \hat{\mathbf{g}}_2, \hat{\mathbf{g}}_3)$ for conjugate tooth flanks with $\hat{\mathbf{g}}_1$ as the meshing line (see Fig. 4).

5.1 The Frenet Frame of the Axodes

Upon gear meshing, the *Frenet frame* $(\hat{\mathbf{f}}_1, \hat{\mathbf{f}}_2, \hat{\mathbf{f}}_3)$ of the axodes with $\hat{\mathbf{f}}_1 = \hat{\mathbf{p}}_{32}$ remains fixed in the gear frame Σ_1 . The second axis $\hat{\mathbf{f}}_2$ equals the spear $\hat{\mathbf{e}}_3$ along the common perpendicular of the gear axes $\hat{\mathbf{p}}_{21}$ and $\hat{\mathbf{p}}_{31}$. In terms of the basis $(\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3)$ we obtain from (3) (see Fig. 1)

$$\begin{pmatrix} \hat{\mathbf{f}}_1 \\ \hat{\mathbf{f}}_2 \\ \hat{\mathbf{f}}_3 \end{pmatrix} = \begin{pmatrix} \cos \hat{\varphi} & \sin \hat{\varphi} & 0 \\ 0 & 0 & 1 \\ \sin \hat{\varphi} & -\cos \hat{\varphi} & 0 \end{pmatrix} \begin{pmatrix} \hat{\mathbf{e}}_1 \\ \hat{\mathbf{e}}_2 \\ \hat{\mathbf{e}}_3 \end{pmatrix}. \quad (24)$$

The origin of this Frenet frame is the striction point $S = (0, 0, \varphi_0)$ of the axodes, the point of intersection between the ISA $\hat{\mathbf{p}}_{32}$ and the common normal of $\hat{\mathbf{p}}_{21}$ and $\hat{\mathbf{p}}_{31}$. The movement of this frame along the axode $\Pi_2 \subset \Sigma_2$ is the rotation Σ_1/Σ_2 about the axis $\hat{\mathbf{p}}_{21}$ with the angular velocity $-\omega_{21}$. Therefore

$$\hat{\mathbf{p}}_{21} = \cos(\hat{\varphi} + \hat{\alpha})\hat{\mathbf{f}}_1 + \sin(\hat{\varphi} + \hat{\alpha})\hat{\mathbf{f}}_3$$

is the permanent Disteli axis of Π_2 . Due to (1), the corresponding Frenet equations (note $\hat{\mathbf{e}}_3 = \hat{\mathbf{f}}_2$) begin with

$$\begin{aligned} \dot{\hat{\mathbf{f}}}_1 &= -\omega_{21}\hat{\mathbf{p}}_{21} \times \hat{\mathbf{f}}_1 = -\omega_{21}\sin(\hat{\varphi} + \hat{\alpha})\hat{\mathbf{f}}_2 \\ &= -\omega_{21}[\sin(\hat{\varphi} + \hat{\alpha}) + \varepsilon(\varphi_0 + \alpha_0)\cos(\hat{\varphi} + \hat{\alpha})]\hat{\mathbf{f}}_2, \end{aligned}$$

which implies for the axode Π_2 the distribution parameter⁴

$$\delta_2 = (\varphi_0 + \alpha_0)\cot(\varphi + \alpha)$$

and the coefficients

$$\hat{\lambda}_2 = -\omega_{21}\sin(\hat{\varphi} + \hat{\alpha}), \quad \hat{\mu}_2 = -\omega_{21}\cos(\hat{\varphi} + \hat{\alpha}).$$

⁴ For the generators of a one-sheet hyperboloid of revolution with semi-axes a, b the absolute value of the distribution parameter equals the secondary semi-axis, i.e., $|\delta| = b$.

The last equation follows from the third Frenet equation $\dot{\hat{\mathbf{f}}}_3 = -\omega_{21}\hat{\mathbf{p}}_{21} \times \hat{\mathbf{f}}_3$ in (13), and it confirms for the dual angle $\hat{\gamma}_2$ between the generator $\hat{\mathbf{p}}_{32} = \hat{\mathbf{f}}_1$ and the Disteli axis $\hat{\mathbf{p}}_{21}$ by (21) $\hat{\gamma}_2 = \hat{\varphi} + \hat{\alpha}$ with $\cot \hat{\gamma}_2 = \hat{\mu}_2/\hat{\lambda}_2$ as *dual curvature* of Π_2 according to [1, Theorem 3].

In a similar way we obtain for Π_3 the distribution parameter

$$\delta_3 = (\varphi_0 - \alpha_0)\cot(\varphi - \alpha)$$

and the coefficients

$$\hat{\lambda}_3 = -\omega_{31}\sin(\hat{\varphi} - \hat{\alpha}), \quad \hat{\mu}_3 = -\omega_{31}\cos(\hat{\varphi} - \hat{\alpha}).$$

The equation $\delta_2 = \delta_3$, which can also be concluded from (5), guarantees the contact between Π_2 and Π_3 at all points of the ISA $\hat{\mathbf{p}}_{32}$.

In the Frenet equations of the auxiliary surface $\Pi_4 \subset \Sigma_4$ with axis

$$\hat{\mathbf{p}}_{41} = \cos(\hat{\varphi} - \hat{\beta})\hat{\mathbf{f}}_1 + \sin(\hat{\varphi} - \hat{\beta})\hat{\mathbf{f}}_3$$

and dual velocity $-\hat{\omega}_{41}$ we obtain the coefficients

$$\begin{aligned} \hat{\lambda}_4 &= -\hat{\omega}_{41}\sin(\hat{\varphi} - \hat{\beta}), \\ \hat{\mu}_4 &= -\hat{\omega}_{41}\cos(\hat{\varphi} - \hat{\beta}). \end{aligned} \quad (25)$$

As a consequence, Π_4 has, by virtue of (19), the distribution parameter

$$\delta_4 = h_{41} + (\varphi_0 - \beta_0)\cot(\varphi - \beta).$$

The equation $\delta_4 = \delta_3 = \delta_2$ can be verified using Eqs. (4), (9), and (12). The axis of Π_4 makes, with all generators Π_4 , the dual angle $\hat{\gamma}_4 = \hat{\varphi} - \hat{\beta}$.

5.2 The Frenet Frame of the Tooth Flanks

According to Theorem 2, any line $\hat{\mathbf{g}}$ attached to the auxiliary surface Π_4 traces conjugate tooth flanks Φ_2 and Φ_3 under the respective relative motions Σ_4/Σ_2 and Σ_4/Σ_3 with the auxiliary surface Π_4 roll-sliding on the axodes Π_2 and Π_3 , respectively. The motion Σ_4/Σ_2 is the composition of Σ_4/Σ_1 with the Frenet motion Σ_1/Σ_2 along Π_2 .

When the dual angle $\hat{\eta}_{\Phi_i}^*$ characterizes the instant Disteli axis of Φ_i we can verify the spatial Euler-Savary equation (see [1])

$$(\cot \hat{\eta}_{\Phi_2}^* - \cot \hat{\eta}) \sin \hat{\xi} = \cot \hat{\gamma}_2 - \cot \hat{\gamma}_4 \quad (30)$$

$$= \cot(\hat{\varphi} + \hat{\alpha}) - \cot(\hat{\varphi} - \hat{\beta})$$

for the motion Σ_4/Σ_2 , which generates Φ_2 .

In the same way we can confirm that the Disteli axis $\hat{\mathbf{g}}_{\Phi_3}^*$ of Φ_3 satisfies

$$(\cot \hat{\eta}_{\Phi_3}^* - \cot \hat{\eta}) \sin \hat{\xi} = \cot \hat{\gamma}_3 - \cot \hat{\gamma}_4$$

$$= \cot(\hat{\varphi} - \hat{\alpha}) - \cot(\hat{\varphi} - \hat{\beta}).$$

Upon subtraction of the two Euler-Savary equations we obtain

$$(\cot \hat{\eta}_{\Phi_2}^* - \cot \hat{\eta}_{\Phi_3}^*) \sin \hat{\xi} = \cot \hat{\gamma}_2 - \cot \hat{\gamma}_3,$$

thereby proving the spatial version of a result which is well known in planar and spherical kinematics, namely

Theorem 5. *Let Φ_2 and Φ_3 be conjugate ruled tooth flanks with permanent line contact. Then the Disteli axes $\hat{\mathbf{g}}_{\Phi_2}^*$ and $\hat{\mathbf{g}}_{\Phi_3}^*$ of the instant meshing line satisfy the Euler-Savary equation for the relative motion Σ_3/Σ_2 between the two gears.*

6. A SPATIAL ANALOGUE OF INVOLUTE GEARING

In planar cycloid gearing there are two auxiliary curves, namely two circles, which usually are laid out in a symmetric relative position with respect to the pole tangent. The same is true on the sphere. However, when the auxiliary circles are specified as great circles they become identical, coinciding with the spherical pole tangent t . The axis $\hat{\mathbf{p}}_{41}$ of the great circle t is orthogonal to the ISA $\hat{\mathbf{p}}_{32}$. The corresponding profiles are involutes of the polodes; they are characterized by the constant pressure angle $\alpha = 0^\circ$.

This is the particular case of involute gearing where the pitch circles coincide with the base circles. These profiles are **not** geometrically feasible because of one reason: At the meshing point P on the instant pole tangent t the profiles have either

- a G^2 -contact with mutual penetration, or
- a cusp, and at external gears the curves open towards opposite sides.

We obtain the corresponding spatial version when we specify the axis $\hat{\mathbf{p}}_{41}$ orthogonal to the ISA $\hat{\mathbf{p}}_{32}$ on the Plücker conoid (see Fig. 1). This is the case we analyze below.

Due to Eqs. (4)–(9), the representation $\hat{\mathbf{p}}_{41} = -\sin \hat{\beta} \hat{\mathbf{e}}_1 + \cos \hat{\beta} \hat{\mathbf{e}}_2$ implies

$$\beta = \varphi + \frac{\pi}{2},^5 \quad \beta_0 = -\varphi_0,$$

$$\hat{\varphi} - \hat{\beta} = -\frac{\pi}{2} + 2\varepsilon\varphi_0. \quad (31)$$

Therefore,

$$\sin(\hat{\varphi} - \hat{\beta}) = -1, \quad \cos(\hat{\varphi} - \hat{\beta}) = 2\varepsilon\varphi_0. \quad (32)$$

From Eqs. (10), (5), and (12) follows for our particular choice

$$\hat{\omega}_{41} = -\omega_{21} \sin(\hat{\varphi} + \hat{\alpha})$$

$$h_{41} = R(\cos 2\alpha + \cos 2\varphi). \quad (33)$$

The auxiliary surface Π_4 is a skew orthogonal helicoid with axis $\hat{\mathbf{p}}_{41}$ and pitch h_{41} , the ISA $\hat{\mathbf{p}}_{32}$ being its initial generator. The invariants of Π_4 are, by virtue of (25),

$$\hat{\lambda}_4 = \hat{\omega}_{41}, \quad \hat{\mu}_4 = -2\varepsilon\varphi_0 \hat{\omega}_{41}. \quad (34)$$

The dual angle between the generators of Π_4 and its axis is

$$\hat{\gamma}_4 = \hat{\varphi} - \hat{\beta} = -\frac{\pi}{2} + 2\varepsilon\varphi_0 \text{ with}$$

$$\cot \hat{\gamma}_4 = \hat{\mu}_4 / \hat{\lambda}_4 = -2\varepsilon\varphi_0.$$

From (4), the distance γ_{40} between axis and generators vanishes if and only if $\varphi = 0$, i.e., $\omega_{21} = -\omega_{31}$.

The generating motions Σ_4/Σ_2 and Σ_4/Σ_3 of the tooth flanks Φ_2 and Φ_3 have the twists $\hat{\mathbf{q}}_{42} =$

⁵One could also set $\beta = \varphi - \pi/2$. However, this has no effect on the auxiliary surface. It only reverses the orientation of $\hat{\mathbf{p}}_{41}$ and changes therefore the sign of ω_{41} and ω_{410} .

$\widehat{\omega}_{42}\widehat{\mathbf{f}}_1$ and $\widehat{\mathbf{q}}_{43} = \widehat{\omega}_{43}\widehat{\mathbf{f}}_1$, respectively; in our particular case we have

$$\begin{aligned}\widehat{\omega}_{42} &= -\omega_{21} [\cos(\varphi + \alpha) + \varepsilon(\varphi_0 - \alpha_0) \sin(\varphi + \alpha)], \\ \widehat{\omega}_{43} &= -\omega_{31} [\cos(\varphi - \alpha) + \varepsilon(\varphi_0 + \alpha_0) \sin(\varphi - \alpha)].\end{aligned}\quad (35)$$

Hence,

$$\begin{aligned}\widehat{\omega}_{43} : \widehat{\omega}_{42} &= \tan(\varphi + \alpha) : \tan(\varphi - \alpha) \\ &= (\varphi_0 + \alpha_0) : (\varphi_0 - \alpha_0).\end{aligned}\quad (36)$$

6.1 The ISA as a Line of Regression

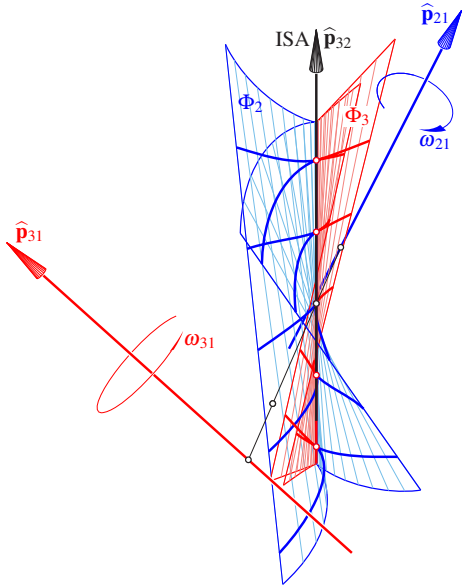


Figure 5: When the ISA coincides with the meshing line $\widehat{\mathbf{g}}$, the singular lines of the two flanks Φ_2 , Φ_3 come together sharing the tangent plane at each point of $\widehat{\mathbf{g}}$; but the flanks open toward opposite sides. The fat red and blue lines indicate sections orthogonal to the ISA.

Analogue to the planar and spherical cases, in spatial cycloid gearing the ISA $\widehat{\mathbf{p}}_{32}$ is a singular generator of the two tooth flanks Φ_2 and Φ_3 . As pointed out in [6, Theorem 5], all its points are uniplanar, the tangent planes along $\widehat{\mathbf{p}}_{32}$ being

distributed just as along a regular generator with distribution parameter $\delta = R \cos 2\alpha$. Figure 8 in [6] reveals that the ISA doesn't look singular at all; it is the common border line of the two components, originating from two symmetrically placed auxiliary surfaces. However, in our particular case the two auxiliary surfaces coincide with the skew helicoid Π_4 . The ISA is, in fact, a line of regression for both tooth flanks. In external gears, as depicted in Fig. 5, the two flanks open toward opposite sides. Hence, when the ISA becomes the meshing line, no transmission of forces can take place. Figure 5 shows the conjugate tooth flanks as wire-frames; the depicted fat red and blue lines being the intersections of the flanks with planes perpendicular to the ISA.

6.2 There is a G^2 -contact at the Striction Point

What corresponds in skew gears to the osculation of tooth profiles when the pole tangent serves as auxiliary curve?

Figure 6 shows an example⁶ where the initial meshing line $\widehat{\mathbf{g}}$ differs from the ISA. But $\widehat{\mathbf{g}}$ is parallel to the ISA and intersects the central tangent of the axodes at right angles. This central tangent passes through the striction point S of the axodes and is parallel to the axis $\widehat{\mathbf{p}}_{41}$ of the auxiliary surface Π_4 (note $\widehat{\mathbf{f}}_3$ in Fig. 1).

The spatial Euler-Savary equation (30) (see [1, Theorem 6]) for the motion Σ_4/Σ_2

$$(\cot \widehat{\eta}^* - \cot \widehat{\eta}) \sin \widehat{\xi} = \frac{\widehat{\omega}}{\lambda} = \cot \widehat{\gamma}_2 - \cot \widehat{\gamma}_4,$$

holds only under $\sin \widehat{\xi} \neq 0$, but we can replace it by the equation [1, page 13]

$$\begin{aligned}\widehat{\lambda} \sin \widehat{\xi} (\cos \widehat{\eta} \sin \widehat{\eta}^* - \sin \widehat{\eta} \cos \widehat{\eta}^*) \\ + \widehat{\omega} \sin \widehat{\eta} \sin \widehat{\eta}^* = 0.\end{aligned}$$

Under the relation $\sin \widehat{\xi} = 0$ (i.e., $\widehat{\mathbf{k}} = \widehat{\mathbf{f}}_3$ in Fig. 4) it is apparent that $\sin \widehat{\eta} \neq 0$ implies $\sin \widehat{\eta}^* = 0$. In other words, when $\widehat{\mathbf{g}} \neq \widehat{\mathbf{p}}_{32}$ intersects the striction tangent $\widehat{\mathbf{f}}_3$ of the axodes at right angles, the

⁶ Data: $2\alpha = 60.0^\circ$, $2\alpha_0 = 70.0\text{mm}$, $\omega_{31} : \omega_{21} = -2 : 3$, and distance between the ISA and the initial meshing line $\widehat{\mathbf{g}}$: $SS_g = 35.0\text{mm}$.

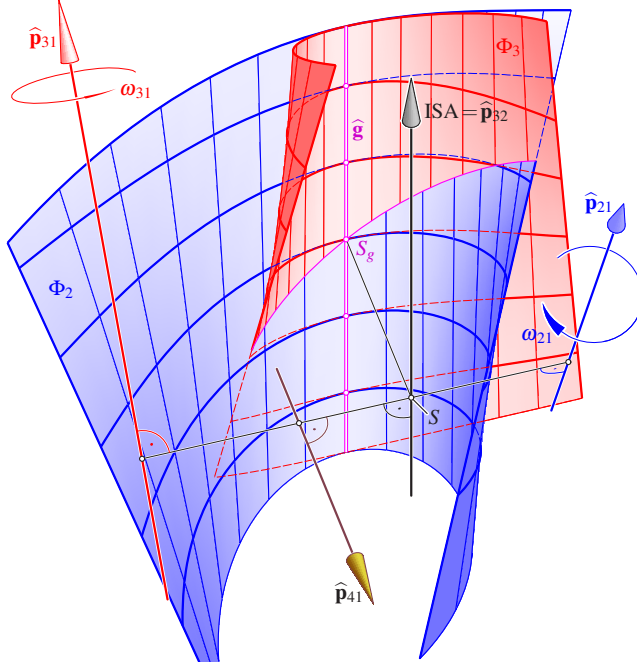


Figure 6: Two conjugate flanks Φ_2 and Φ_3 with G^2 -contact at the common striction point S_g . The meshing line $\hat{\mathbf{g}}$ is parallel to the ISA and a cylindric generator of Φ_2 and Φ_3 .

Disteli axis $\hat{\mathbf{g}}^*$ coincides with the ISA. The same holds for the motion Σ_4/Σ_3 , which means that under this condition the two tooth flanks share the instant Disteli axis. According to Theorem 4, Φ_2 and Φ_3 have a G^2 -contact at the common striction point S_g .

In Fig. 6, the fat blue and read curves, which are in contact at marked points on the meshing line $\hat{\mathbf{g}}$, are level lines of the two flanks, i.e., intersections with planes orthogonal to the ISA. The mean section shows the G^2 -contact at the striction point S_g , which causes the penetration.

The case of osculating cylindrical or spherical tooth flanks is misleading. In the true spatial version there is no G^2 -contact at all other points of $\hat{\mathbf{g}}$ for one reason: According to Theorem 4, in this case the condition $\dot{\tilde{\delta}}(0) = c \dot{\delta}(0)$ must be

satisfied. However, because of the permanent line contact the flanks have the same distribution parameter $\tilde{\delta}(t) = \delta(t)$ for each $t \in I$. This implies $\tilde{\delta}(0) = \dot{\delta}(0)$, but by Eqs. (28), (29) and (36), the constant c with $\hat{\lambda}_{\Phi_3} = c \hat{\lambda}_{\Phi_2}$ is

$$\begin{aligned} c &= \tan(\varphi + \alpha) / \tan(\varphi - \alpha) \\ &= (\varphi_0 + \alpha_0) : (\varphi_0 - \alpha_0) \neq 1. \end{aligned}$$

The different poses depicted in Fig. 7 reveal that there is also a mutual penetration of the conjugate tooth flanks Φ_2 and Φ_3 at the other poses. Since the surfaces share this curve of intersection as well as the tangent planes at all points of the meshing line, there must be a G^2 -contact at the point where the curve of intersection meets the meshing line. This point is close to the mar-

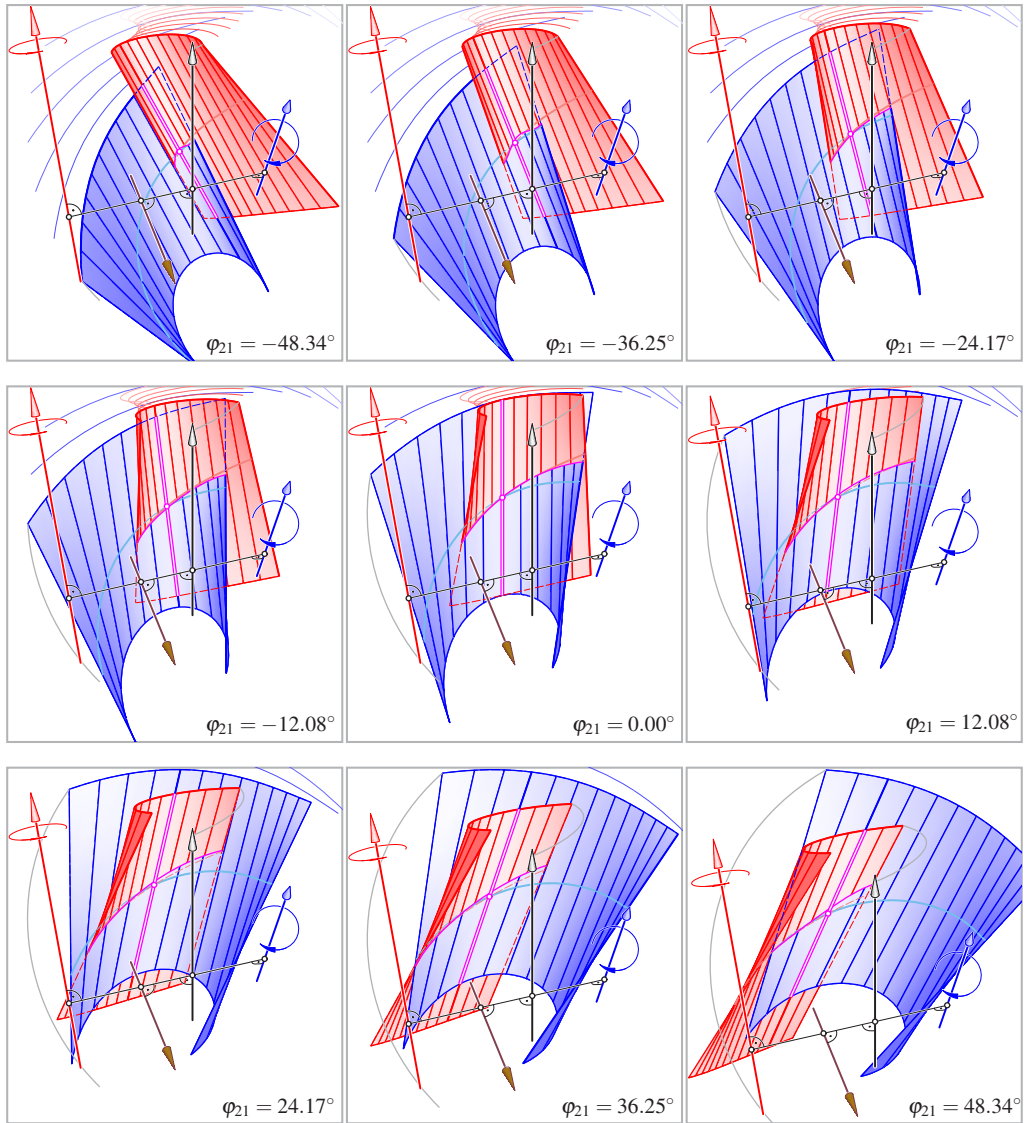


Figure 7: Snapshots of the penetrating tooth flanks with their striction curves upon meshing.

ked striction point; however it can be proved that the point of G^2 -contact must be different from the striction point S_g up to the symmetric case $\omega_{31} = -\omega_{21}$; hence $\varphi_0 = 0$.

6.3 A Spatial Analogue of Octoidal Gears

In the plane as well as on the sphere, the *generalized Camus Theorem* states that for any curve c_4 attached to the auxiliary curve $p_4 \subset \Sigma_4$ the envelopes c_2 and c_3 under motions Σ_4/Σ_2 and Σ_4/Σ_3 , respectively, are conjugate tooth profiles.

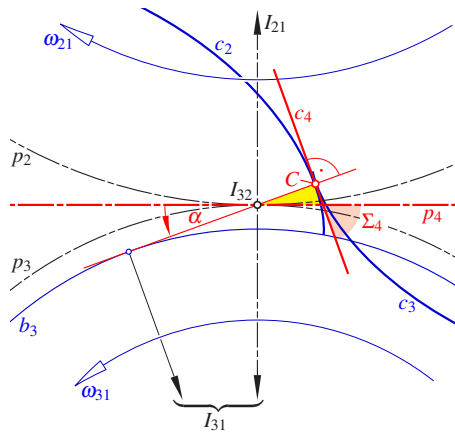


Figure 8: Planar version of the generalized Camus Theorem in the particular case leading to involute gears.

In the particular planar case, depicted in Fig. 8, the auxiliary curve p_4 is the pole tangent t and the attached curve c_4 is a line. In all its poses, the line c_4 shows the same inclination with respect to the gear frame Σ_1 . At each pose, the enveloping point C of c_4 is the pedal point with respect to the instant pole I_{32} . The right-angled triangle enclosed by c_4 , p_4 and the line $I_{32}C$ shows that the pressure angle α remains constant, which leads to the case of involute gearing.

The foregoing statement does not hold in spherical geometry since for spherical triangles the sum of the interior angles is not constant. This sum is always greater than 180° , the amount

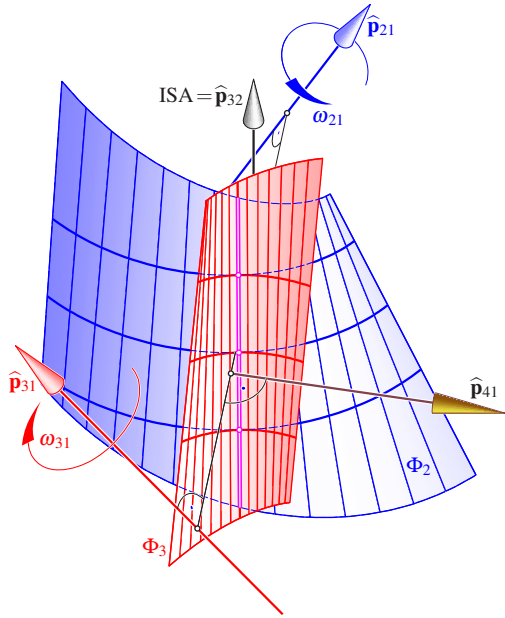


Figure 9: Skew gears with torsors as conjugate tooth flanks Φ_2, Φ_3 and permanent line contact. The fat red and blue lines indicate sections orthogonal to the meshing line.

by which the sum exceeds 180° being proportional to the area of the triangle. Therefore, we cannot conclude for the analog specification in bevel gears that the pressure angle α is constant; it increases with the distance between I_{32} and c_4 . We obtain what is known as *octoidal gears*, as reported in [3].

The validity of the spatial analogue of the generalized Camus Theorem was proved in [6, Theorem 6]: For any surface Φ_4 attached to the auxiliary surface Π_4 the envelopes Φ_2 and Φ_3 under the respective relative motions Σ_4/Σ_2 and Σ_4/Σ_3 are conjugate tooth flanks. We choose again Π_4 as the skew orthogonal helicoid and specify Φ_4 as a plane. Then we obtain a pair of conjugate torsors Φ_2, Φ_3 with permanent line contact. In Fig. 9 one example is depicted which indicates

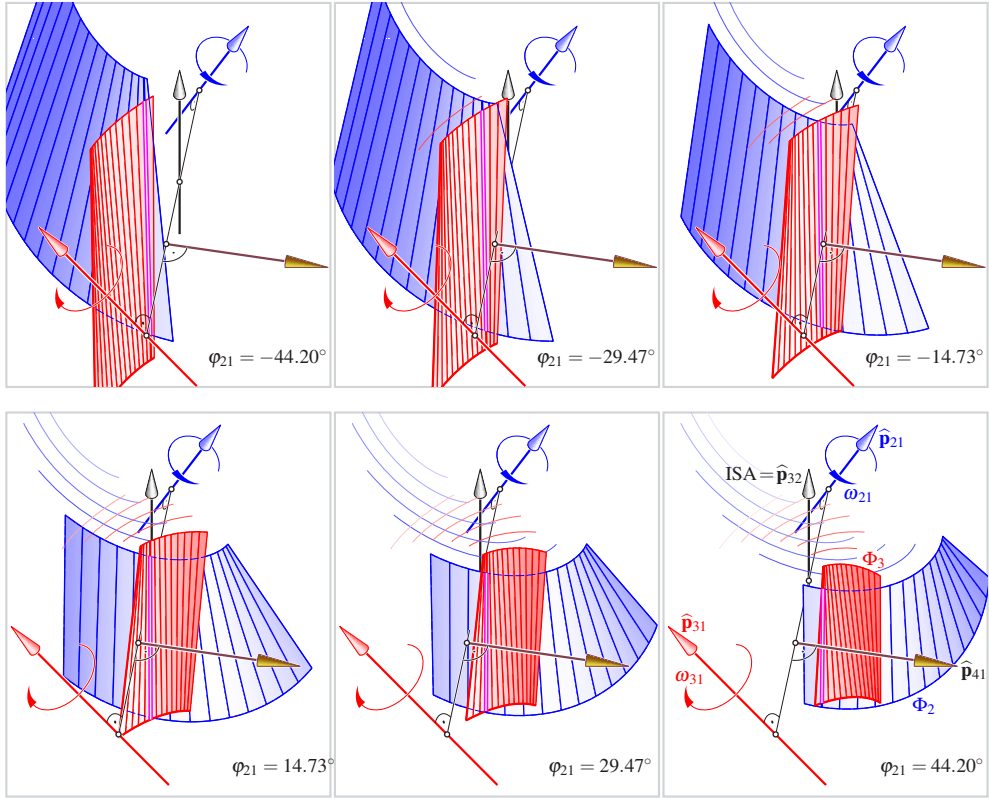


Figure 10: Snapshots of the conjugate torses Φ_2 and Φ_3 upon meshing ($\omega_{31} : \omega_{21} = -2 : 1$).

that these flanks should work correctly. Contrary to the general case of J. Phillips' involute gearing [7], contact is not punctual, but along a line.

The fat red and blue curves in Fig. 9 are the intersections of the flanks with planes perpendicular to the instant meshing line, which is depicted as magenta double line. Figure 10 shows snapshots of the conjugate torses upon meshing.

7. CONCLUSIONS

Based on the Camus Theorem and on Martin Disteli's work, we showed in this paper that the flanks of spatial cycloid gears can be synthesized by means of an auxiliary surface. Upon choos-

ing the skew orthogonal helicoid as auxiliary surface, the tooth flanks of the spatial equivalent of octoidal gears are obtained. The final example with torses as conjugate tooth flanks looks promising but still needs a detailed analysis.

REFERENCES

- [1] Stachel, H. (2000). Instantaneous spatial kinematics and the invariants of the axodes, *Proc. Ball 2000 Symposium*, Cambridge (article no. 23).
- [2] Blaschke, W. (1960). *Kinematik und Quaternionen*, VEB Deutscher Verlag der Wissenschaften, Berlin.

- [3] Figliolini, G., Angeles, J. (2005). Algorithms for Involute and Octoidal Bevel-Gear Generation *ASME J. Mech. Design* 127(4), 664–672.
- [4] Figliolini, G., Stachel, H., Angeles, J. (2007). A new look at the Ball-Disteli diagram and its relevance to spatial gearing, *Mech. Mach. Theory* 42(10), 1362–1375.
- [5] Figliolini G., Stachel H., Angeles J. (2013). On the synthesis of spatial cycloidal gears, *Meccanica* 48(5), 1239–1249.
- [6] Figliolini, G., Stachel, H., Angeles, J. (2013). On Martin Disteli's spatial cycloidal gearing, *Mech. Mach. Theory* 60(1), 73–89.
- [7] Phillips, J. (2003). *General Spatial Involute Gearing*, Springer Verlag, New York.
- 2. Hellmuth Stachel, Ph.D, is Professor emeritus of Geometry at the Institute of Discrete Mathematics and Geometry, Vienna University of Technology, Doctor h.c. University of Dresden, and editor in chief of the "*Journal for Geometry and Graphics*". His research interests are in Higher Geometry and Kinematics. He can be reached by e-mail: stachel@dmg.tuwien.ac.at, by fax: +43-1-58801-10493, or through the postal address: Institute of Discrete Mathematics and Geometry, Vienna University of Technology, Wiedner Hauptstraße 8–10/104, A-1040 Wien, Austria.
- 3. Jorge Angeles is James McGill Professor of ME, McGill University, Robotic Mechanical Systems Laboratory's founder and Director. Author/co-author of various books in kinematics and dynamics of mechanical systems, numerous publications in research journals and conference proceedings. Three international patents. Fellow of The Royal Society of Canada; one of the 12 Honorary Members of IFToMM; Fellow of CSME, ASME and IEEE, Fellow of the Canadian Academy of Engineering. Doctor h.c. University of Guanajuato, Mexico. He can be reached by e-mail: angeles@cim.mcgill.ca.

ABOUT THE AUTHORS

1. Giorgio Figliolini is Associate Professor of Mechanics of Machinery at the University of Cassino & Southern Lazio, Italy, in which he got the Master Degree in Mechanical Engineering with laude. In 1998, he was Visiting Scholar at the CIM of the McGill University in Montreal (Canada). His main research interests are in the fields of Kinematics, Mechatronics and Robotics. In 2014, he got the Italian National Scientific Habilitation as Full Professor of Mechanics of Machinery. He can be reached by e-mail: figliolini@unicas.it, by fax: +39-0776-2993704, or via the postal address: Dept. of Civil & Mechanical Engineering (DICEM), University of Cassino & Southern Lazio, Via G. Di Biasio 43, 03043 Cassino (FR), Italy.

A STUDY ON THE PARAMETRIC ARCHITECTURE - DEVELOPMENT OF COMPUTER GRAPHIC MATERIALS FOR ARCHITECTURAL DESIGN EDUCATION -

Naomi ANDO¹, Shota ISHII¹, Nobuhiro YAMAHATA² and Akihiro SHIBATA³

¹Hosei University, Japan ²Tohoku University of Art and Design ³Kagoshima University

ABSTRACT: This study develops computer graphic materials to learn about abstract features of architectural forms. First, we created 3D models that simplify distinctive architectural forms. In this step, the details are omitted and much attention is paid to simple modeling of the volume (form of architecture) and the elements (such as openings, roofs, columns) of the structure. Second, we defined the characteristics of the volume and elements using parameters and 3D models, in which the characteristics of architectural forms are variable by the parameters. Through these steps, the 3D models are able to describe existing architectures using parameters, while they describe non-existent imaginary architectural forms by different parameters. In this study, we developed a teaching tool that enables beginner architecture students learn the characteristics of architectural forms by operating 3D models on a computer by changing parameters.

Keywords: Architecture, Computer Graphics, Algorithmic Design, Education.

1.INTRODUCTION

Architectural forms have various parameters such as width, height, depth, number of stories, shape of roof, and size or arrangement of elements such as windows, entrances, and columns. Architectural forms are generally designed according to the principles of structure and function. However, for the public, the form would in many cases become established as the image of architecture. For beginning architec-

tural students, it is considered important to know the various possible forms of architecture along with learning the structures and functions.

In this paper, we attempted to develop computer graphic materials to help learn abstract features of architectural forms. Table 1 shows a list of “Parametric Architecture” as computer graphic materials, presented in this paper.

Table 1: List of parametric architectures.

Architecture	Location	Year	books / URL for reference
1 Eiffel Tower	Paris, France	1889	The Eiffel Tower, Taschen America, 2008
2 Parthenon	Athens, Greece	BC 6C	Metron Ariston, Aeropos publications, 2003
3 Gateway Arch	St. Louis, U.S.A.	1965	http://www.gatewayarch.com/
4 Restaurant in Xochimilco	Mexico City, Mexico	1958	P.L.Nervi, F.Candela, Bijyutu Shuppan-Sha, 1970
5 Palazzetto dello Sport Rome	Rome, Italy	1957	P.L.Nervi, F.Candela, Bijyutu Shuppan-Sha, 1970
6 Church of Saint Francis of Assisi	Pampulha, Brazil	1943	Oscar Niemeyer, Bijyutu Shuppan-Sha, 1969
7 Kyusendo Museum	Kumamoto, Japan	1984	Shinkenichiku 1984-10, Shinkenichiku-Sha
8 Kimbell Art Museum	Fort Worth, U.S.A.	1972	Kimbell Art Museum, Dohosha 1984
9 Tomihiro Art Museum	Gunma, Japan	2005	Shinkenichiku 2005-04, Shinkenichiku-Sha
10 Habitat '67	Montreal, Canada	1967	http://www.habitat67.com/
11 Sendai Mediatheque	Sendai, Japan	2000	Sendai Mediatheque, GA Detail No.2, 2001
12 Zollverein School	Essen, Germany	2006	Shinkenichiku 2006-11, Shinkenichiku-Sha

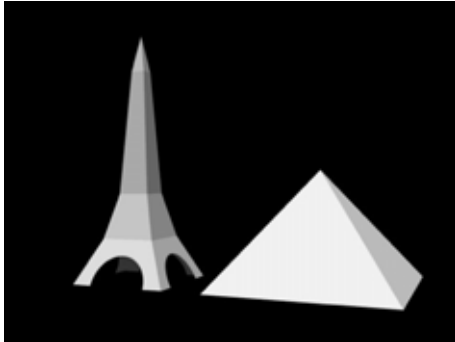


Figure 1: Eiffel Tower and Great Pyramid.

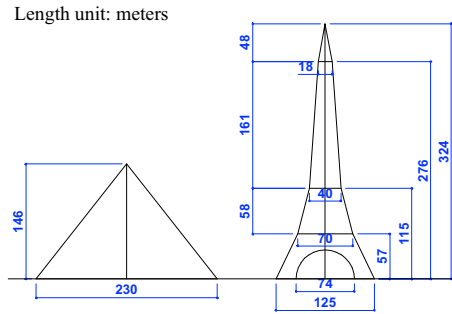


Figure 2: Elevations (Great Pyramid and Eiffel Tower).



Figure 3: Eiffel Tower.



Figure 4: Great Pyramid of Khufu.

2.EIFFEL TOWER

As a basic example, simplified models and elevations of the Eiffel Tower (1889) in Paris, France and the Great Pyramid of Khufu (26th century BC) in Giza, Egypt, are shown in Figures 1 and 2. They are gigantic structures constructed with steel frames and stone masonry, respectively. Photographs are shown in Figures 3 and 4.

The form of the Eiffel Tower is composed of three truncated pyramids; the truncated pyramid at the bottom has arches and one quadrangular pyramid on the top.

The form of the Eiffel Tower is generated using the following procedure:

- 1) Place the first truncated pyramid at the bottom. The side length is 125m at the bottom and 70m on the top. The height is 57m.

- 2) Place the second truncated pyramid. The side length on the top is 40m. The height is 58m.
- 3) Place the third truncated pyramid. The side length on the top is 40m. The height is 58m.
- 4) Place the pyramid on the top. The height is 48m.
- 5) Place two cylinders with a diameter of 74m at the bottom, and hollow out the truncated pyramid.

Regarding the above procedure, when the height of the three truncated pyramids, and the radius of the cylinders are set to zero; the side length and the height of the top pyramid are set to 230m and 146m, the Great Pyramid emerges. In this way, the forms are generated by an algorithm with parameters that describe their compositions and dimensions.



Figure 5: Parthenon.

For students, it is essential to study the dimensions and characteristics of architectural forms by creating drawings and models. A study that helps generate architectural forms accurately by finding adequate parameters, will assist in learning about architectural forms. Moreover, deforming architectural forms with parameters may be useful in imagining new possible designs.

3.PARTHENON

Figures 5 and 6 indicate a photograph and drawings, which are simplified representations of the dimensions and the composition of the Parthenon (5th century BC), an example of ancient Greek architecture in Athens, Greece. In this form, 8 cylindrical columns stand on the foundation stone from east to west, and 17 col-

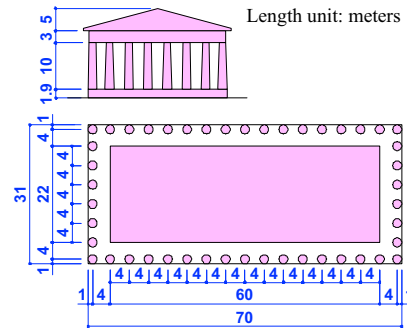


Figure 6: Plan and Elevation (Parthenon).

umns stand from north to south, to hold up the pediment.

The simplified model, shown in Figure 7, is created by using Shade 3D [1] (3DCG Application). The form is generated by an algorithmic script coded in Python [2]. The script starts with a description of the parameters as follows:

1. w = 31000 #width
2. d = 70000 #length
3. ww = 21500 #width of inner core
4. dd = 60200 #length of inner core
5. ncw = 6 #columns in cross direction
6. ncd = 15 #columns in long direction
7. hbase = 1900 #height of base
8. hcolumn = 10400 #height of columns
9. hcornice = 2700 #height of entablature
10. hpediment = 5000 #height of pediment

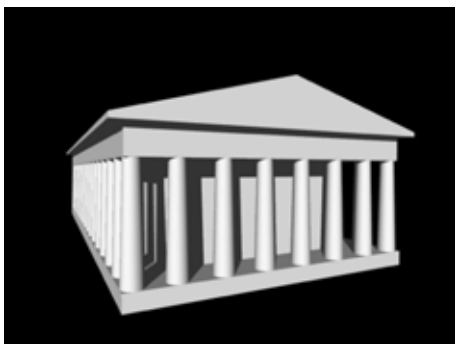


Figure 7: Simplified model (Parthenon).

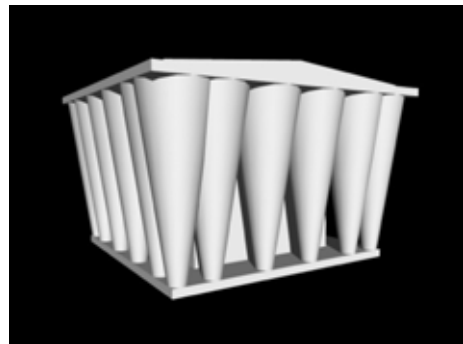


Figure 8: Deformed model (Parthenon).



Figure 9: Gateway Arch (left: Photo, middle: Simplified model, right: Deformed model).

- 11. hpbse = 500 #height of pediment base
- 12. r1 = 1000 #radius of columns at bottom
- 13. r2 = 750 #radius of columns at top
- 14. cl = 1000 #cantilever of entablature

By changing the above parameters, the form will be deformed as exemplified in Figure 8. An exercise, which aims to acquire the correct form by finding the adequate parameters, may be effective for students to learn the composition and the dimensions of architectural forms.

4.GATEWAY ARCH

Gateway Arch: The Jefferson National Expansion Memorial (1965), designed by Eero Saarinen (1910-1961), is an arched structure constructed in St. Louis, U.S.A. Tram cars, which climb up to the observation gallery at the top, are built inside the arch. Both the width and the height of the arch are 192m. The cross-section of the arch has the form of an equilateral triangle; the length of one side is 16.5m at the bottom and 5.2m at the top.

The shape of the arch follows the catenary curve represented by the following equation (1):

$$y=a \cdot \cosh(x/a) \quad (1)$$

Figure 9 shows the photograph, the simplified model, and a deformed model. Besides the width and height of the arch, the shape of the cross-section of the arch is described by polygon with any number of vertex with variable radius.

Similar to this example, there are many ar-

chitectural forms that are represented by formulas. As long as a form is described by a formula, it is possible to control the embodiment by changing some parameters in the formula. By manipulating the form, students will be able to confirm the aesthetic forms made from the formulas.

5.RESTAURANT IN XOCHIMILCO

Los Manatiales Restaurant (1958), designed by Felix Candela (1910-1997), is a restaurant in Xochimilco, Mexico City, Mexico.

Figure 10 shows the simplified model and a deformed model. A beautiful shape of eight petals is modeled by four hyperbolic paraboloid vaults, the ends of which are cut obliquely; four vaults are rotated by a 45 degree angle.

The parameters for the simplified model are as follows:

- 1) Length of vault on the roof top from the center to the tip: 12.5m
- 2) Length of vault at the ground from the center to the tip: 9.1m
- 3) Width / Height of parabolic curve at the center: 2.6m / 3.6m
- 4) Width / Height of parabolic curve at the tip on the ground: 7.5m / 4.8m

In the deformed model, the above parameters are altered.

6.PALAZZETTO DELLO SPORT

The Palazzo dello Sporto (1957), designed by Pier Luigi Nervi (1891-1979), is a gymnasium



Figure 10: Restaurant in Xochimilco (left: Simplified model, right: Deformed model).

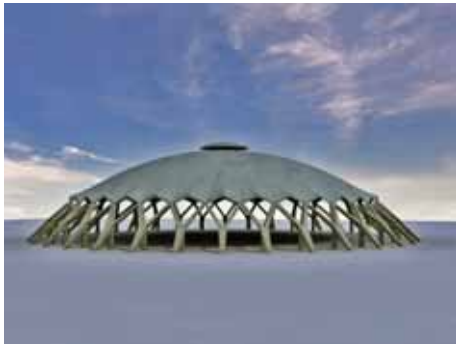


Figure 11: Palazzo dello Sport (left: Simplified model, right: Deformed model).

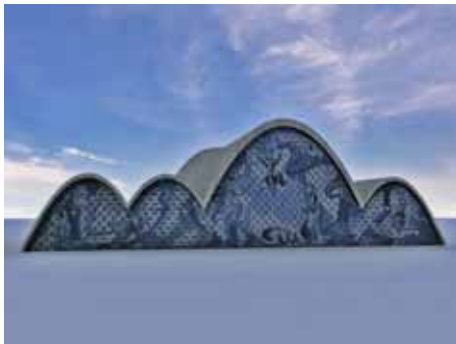


Figure 12: Church of Saint Francis of Assisi (left: Simplified model, right: Deformed model).

constructed in Rome, Italy, for the 1960 Olympic Games. The dome, approximately 60m in diameter, is supported by Y-shaped pre-cast concrete frames.

Figure 11 shows the simplified model and a deformed model. The parameters for the simplified model are as follows:

- 1) Height / Radius of dome: 15m / 7m

2) Ratio of the height of Y-shaped frame and dome: 0.4

3) Number of Y-shaped frames: 36

For the deformed model, the ratio of the height of Y-shaped frame to dome is not changed, while the height of the dome is altered to 20m, and the number of Y-shaped frames is altered to 24.

7.CHURCH OF SAINT FRANCIS OF ASSISI

The Church of Saint Francis of Assisi (1943), designed by Oscar Niemeyer (1907-2012), located by a lake in Panpura, Brazil, has a series of arch-shaped roofs. Beautiful murals in blue paint grace the outer walls.

Four reinforced concrete vaults with a parabolic arch shape are arranged in a wave shape. The largest vault extends while changing the shape in the long direction.

Figure 12 shows the simplified model and a deformed model. The parameters for the simplified model are as follows:

- 1) Referential width / height of vault: 8m / 5m
- 2) Length of vault: 12m
- 3) Length of extended vault: 15m
- 4) Repeat number and scale of vaults
- 5) Position and Scale of extended vault

For the deformed model, the repeat number is set to 8 and scales are set to 1.8, 1.2, 1.5 for the second, fifth, and sixth vaults (other scales are 1.0). The scales of the extended vaults are set to 1.2 and 0.8 for the second and sixth extended vaults.

8.KIMBELL ART MUSEUM

Kimbell Art Museum (1972), designed by Louis Kahn (1901-1974), is a museum in Fort Worth, Texas, U.S.A. The building has vaulted ceilings with slit-shaped skylights.

Figure 13 shows the photograph, the simplified model, and a deformed model. The parameters of the simplified model are as follows:

- 1) Width/Height/Length of vault: 7m/2.4m/30m
- 2) Curvature of vault (tangent at start point): 1.12
- 3) Width between two vaults: 2.4m
- 4) Thickness / Height of wall: 600mm / 3.85m

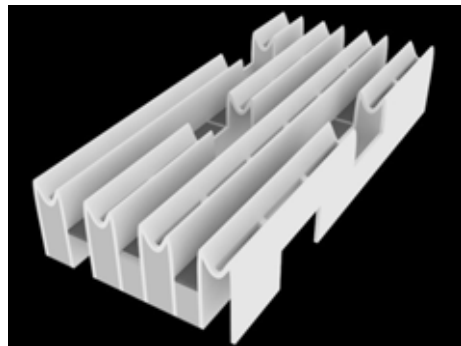


Figure 13: Kimbell Museum and Deformed model.

5) Array of vaults: 6 rows x 3 columns

The deformed model is generated by altering all the above parameters.

9.KYUSENDO MUSEUM

Kyusendo Museum (1984), designed by Yasufumi Kijima (1937-1992), is a museum dedi-



Figure 14: Kyusendo Museum.

cated to trees and forests located in Kumamoto Prefecture, Japan. The structure consists of seven intersecting domes of 7.5m diameter built into the steep slopes of the mountain. Figure 14 shows the photograph, and Figure 15 shows the simplified model and a deformed model.

The algorithm to generate the model is described by Grasshopper [3], which is a plug-in for Rhinoceros [4] (3D modeling application). By using Grasshopper, an algorithm is visually composed by arranging the components on a graphic screen. A screenshot of Grasshopper and Rhinoceros used to generate one form of a dome is shown in Figure 16. The parameters are as follows: radius (R); height (H) of the cylinder below the dome; coordinates (X, Y, Z) of the center of the dome. When these parameters are changed by moving the slider on

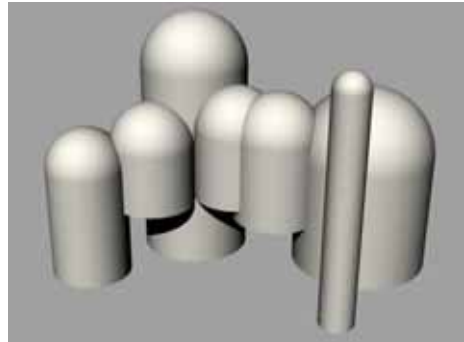
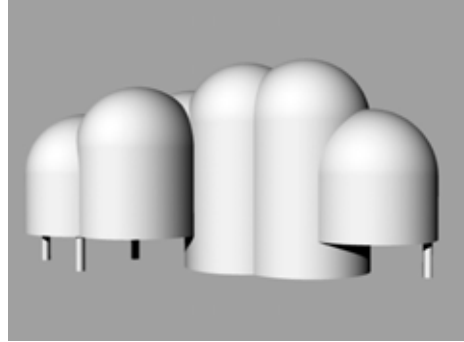


Figure 15: Models (Kyusendo Museum).

Grasshopper, the model on Rhinoceros follows its properties interactively.

10.TOMIHIRO ART MUSEUM

Tomihiro Art Museum (2005), designed by Makoto Yokomizo, located in Gunma Prefecture, Japan, is a single-story museum with 33 circular galleries and rooms. A simplified

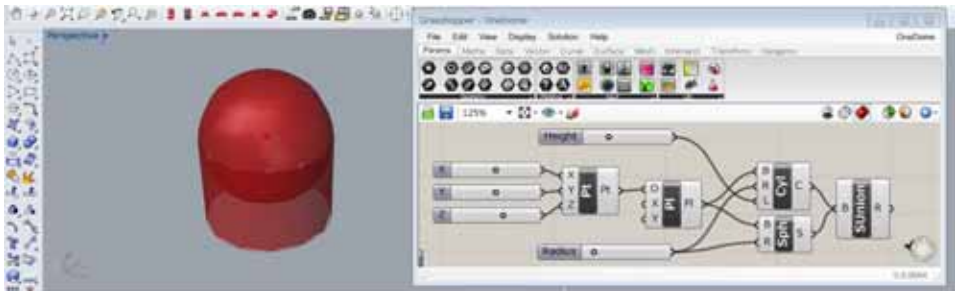


Figure16: Screenshot of Rhinoceros and Grasshopper.

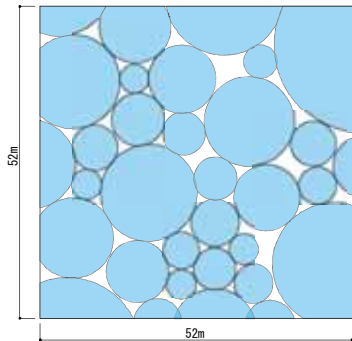


Figure 17: Simplified plan and model (Tomihiko Art Museum).

ground plan and model are shown in Figure 17.

While architects often use geometric shapes, the form is designed to the smallest detail in exemplary architecture. At first glance, the flat surface of the Tomihiko Art Museum looks like a collection of random circles; however, the room configuration is actually based on very detailed architectural planning. It is not easy to describe detailed architectural planning with an algorithm.

On the other hand, creating an algorithm that simply places objects randomly is rather easy. Figure 18 shows a piece of furniture designed by one of the authors [5]. This shelving unit, made by paper tubes, was designed using a script which randomly places different sized circles in a fixed area.



Figure 18: Paper Tube Furniture.

Figure 19 is a model of the same sort of the motif, which has expanded circles into three dimensional spheres. Figure 20 shows a project for museum which uses a sphere motif, de-

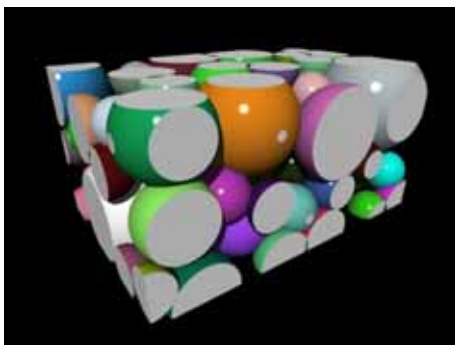


Figure 19: Design by Spheres.



Figure 20: Project for art museum.



Figure 21: Habitat '67.

signed by a graduate student. The final form of this design was decided through detailed architectural planning. However, in the initial stages of the design, sketches are made by utilizing an algorithm with random parameters.

11.HABITAT '67

Habitat '67 (1967), designed by Moshe Safdie, shown in Figure 21, is a housing complex located in Montreal, Canada. 158 houses were composed from 354 individual units. The floor size of each unit is 11.4 x 5.1m (38 x 17ft). Each unit was manufactured in a factory, and stacked up on site to a height of 12 layers.

In spite of its careful design, it appears like randomly stacked boxes. Figure 22 shows a shape created by randomly stacking boxes on a scale similar to Habitat '67. Within this algo-



Figure 22: Deformed model.

rithm, the box size, number, number of layers, and density can be specified in the parameters. Figure 23 shows a model created by changing the parameters.

Figure 24 shows a project for a complex building designed by a graduate student. The form was designed using the stacked box motif; however, as the boxes used have different sizes corresponding to different functions, the design process was complicated. It was difficult to describe the form using a script. Thus, even if parametric scripts are useful to sketch design ideas, the whole design process may not be automatically completed by a script [6].

12.SENDAI MEDIATHEQUE

Sendai Mediateque (2000), designed by Toyo Ito. located in Sendai, Japan, shown in Figure



Figure 23: Deformed model.



Figure 24: Project for complex building.



Figure 25: Sendai Mediatheque.

25, is a 7-story complex containing a library, galleries, conference rooms, etc. Thirteen steel tubes of light forms support the floors with different ceiling heights.

While the thirteen tubes have been designed and constructed elaborately, they resemble plants growing upward in organic shapes. The above shape in Figure 26 shows a model which simulates this form (the fluctuations in the tubes have been randomized instead of following the real design). The figure below shows a deformed model.

13.ZOLLVEREIN SCHOOL

Zollverein School (2006), designed by SANAA (Kazuyo Sejima and Ryuei Nishizawa), is a university building located in Essen, Germany. Figure 27 shows its elevation. Windows of four different sizes are carefully arranged on each side of a 35m approx. cube structure.

Figure 28 is the simplified model based on the elevation. Figure 29 shows a deformed model in which the window placement has been randomized, ignoring the elevation.

Figure 30 shows a model that combines the previous randomly stacked box model, shown in Figure 23, and the randomly placed windows. Due to the feeling of scale that appears in the windows, this form may look like an apartment complex. However, in the actual design, the windows must be arranged with some

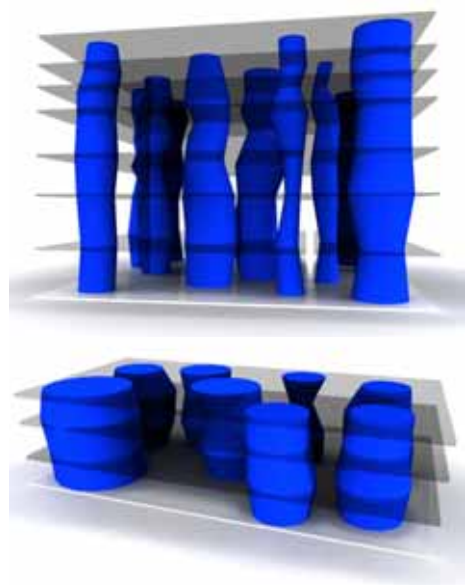


Figure 26: Deformed models.

purpose; it is unlikely that they were arranged randomly.

The elevation models of Notre Dame Cathedral (1345) in Paris, France, is shown in Figure 31. According to these models, based on the actual construction, window-opening ratio (the ratio of the window area in relation to the wall) is 17.5%. Figure 32 displays a project for church designed by a graduate student. In this project, a script which randomly arranges the windows is used; however the window-opening ratio was controlled so that it would be equal to that of Notre Dame Cathedral in Paris. This is an example that uses a script to determine a design of architectural form (in this case, the window-opening ratio), even if the form is randomly generated. When a goal to use the parametric script is specified, it may be possible to utilize “Parametric Architecture” as a means of designing a parametric form. If this is not the case, it should be noted that the manipulation of parameters would merely become a method to

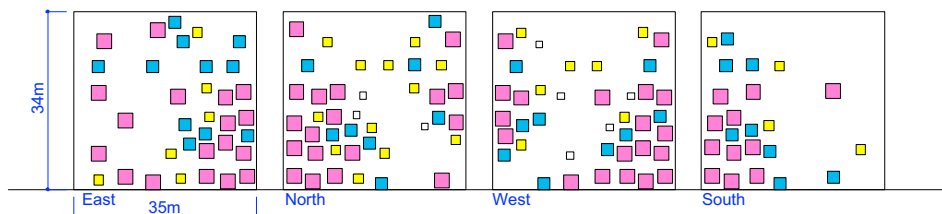


Figure 27: Elevations (Zollverein School).



Figure 28: Simplified model (Zollverein School).

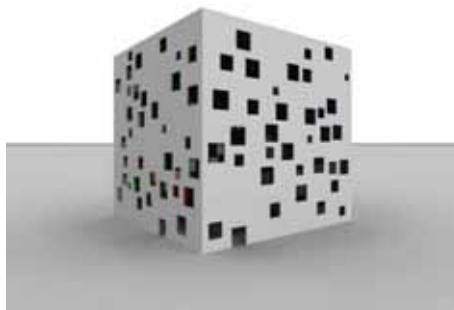


Figure 29: Deformed model.

play around with shapes.

14.CONCLUSION

In this study, we pursued “Parametric Architecture” as a tool for architectural education. Students can operate 3D models on computers according to their interests and learn the scale and composition of architectural forms. Furthermore, using these operational processes, we expect them to gain knowledge about the beau-

ty of renowned architecture and the pleasure of designing one’s own structures.

ACKNOWLEDGMENTS

The authors wish to thank Mr. Takahito Abe, Mr. Atsushi Funaoka, and Mr. Takeshi Tomozawa, former graduate students of Hosei University, for their efforts in completing the projects shown in Figures 20, 24, and 32.

REFERENCES

- [1] e frontier Inc., Shade 3D:
<http://shade3d.jp/>
- [2] Python Software Foundation:
<https://www.python.org/>
- [3] Grasshopper, Algorithmic Modeling for Rhino:
<http://www.grasshopper3d.com/>
- [4] Robert McNeel & Associates., Rhinoceros:
<http://www.rhino3d.com/>
- [5] N. Ando and A. Shibata. Fabrication of Furniture by Using Algorithmic Design. Scientific Bulletin of the "Politehnica" Uni-



Figure 30: Design by Boxes.

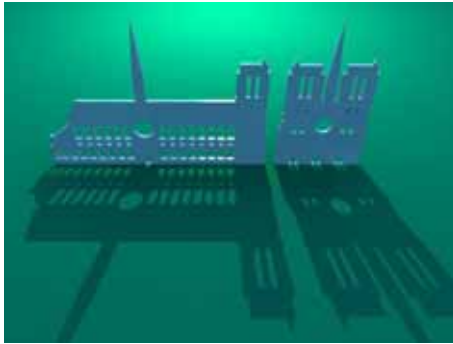


Figure 31: Study on lights.

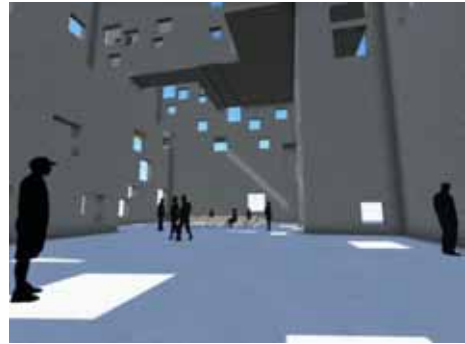


Figure 32: Project for Church of Lights.

versity of Timisoara, Romania, ISSN:1224-6042, pages 79–82. 2013.

- [6] N. Ando, A. Shibata, Y. Nagano and N. Yamahata. Design by Boxes - Study on an Educational Method for Architectural Form Generation, Part 2. Acta Technica Napocensis, Series Applied Mathematics and Mechanics 52, Vol.1b, ISSN:1221-5872, Technical University of Cluj-Napoca, pages 591-596. 2009.

ABOUT THE AUTHORS

1. Naomi ANDO, Dr.Eng, is a professor of Hosei University. His research interests are the analysis of forms and spatial images, application of computer graphics to architectural and urban planning. He can be reached by e-mail: n-ando@hosei.ac.jp, by fax: +81-3-5228-1403, or via postal address: Hosei University / Department of Architecture / Faculty of Engineering and Design, 2-17-1 Fujimi, Chiyoda-ku, Tokyo 102-8160, Japan
2. Shota Ishii received his master's degree in Architecture from Hosei University in 2012. He is currently enrolled in the doctoral course at Department of Architecture, HOSEI University. His research interests are analysis of architectural drawings, design of architecture and graphics. His e-

mail address is as follows: shota.ishi-i0322@gmail.com.

3. Nobuhiro YAMAHATA is a professor of Tohoku University of Art and Design. His research interests are traditional architecture, material re-use and conversion system in architectural design, sustainable design and scenery. He can be reached by e-mail: yamahata.nobuhiro@aga.tuad.ac.jp, by fax: +81-236272252 or through postal address: Tohoku University of Art and Design / Department of Architecture and Environmental Design / Faculty of Design, 3-4-5 Kamisakurada, Yamagata 990-9530, Japan.
4. Akihiro SHIBATA, Dr. Eng, is an architect and also a associate professor of Kagoshima University. His research interests are methods of designing forms, aesthetics of architectures and streetscape. He can be reached by e-mail: shibata@aae.kagoshima-u.ac.jp, by fax: +81-99-285-8311, or through postal address: Kagoshima University / Dept. of Architecture & Architectural Engineering / Faculty of Engineering, 1-21-40, Korimoto, Kagoshima City, Kagoshima, 890-0065, JAPAN.

STUDY ON THREE-DIMENSIONAL VISUALIZATION OF THE CABLE-STAYED BRIDGE

Ziru WANG¹, Chunxia HU¹, Mingqiu Li²

¹Dalian University of Technology, China ²Dalian Ocean University, China

Abstract: Developed on the base of VC++ platform, combined with OpenGL graphic library and the characteristic of cable-stayed bridge, adopting parameterized modeling method, this paper explores to realize the 3-D visualization of cable-stayed bridge. With the characteristics of high efficiency, clear categorizing parameters, convenient operating system, and this method provides a new approach for the comparison of bridge schemes.

Key Words: VC++; OpenGL; Parameterized; Cable-stayed Bridge; Visualization

1. INTRODUCTION

It is significance for bridge of the 3D simulation in modern bridge optimization design and bridge joint construction. Due to bridge a variety of sizes, complex structures and forms and low levels of standardization, it is difficult to make a three-dimensional modeling and visual simulation of the bridge[1]-[2]. At present, the three-dimensional modeling of bridges is usually treated in AutoCAD or 3ds Max manually, but this method has the faults of fussy process, randomness, and can not conducive to accurate modeling. If the model is formed, its geometry shape is difficult to change. In order to overcome these deficiencies to make the bridge modeling universal, this paper develops three-dimensional visualization system of cable-stayed bridge, which is based on WINDOWS operation platform by using MFC of VC++ and OpenGL graphics library[3]-[4], combining the characteristics of cable-stayed bridge, and adopting parameterized method for three-dimensional modeling of cable-stayed bridge. The study has a very important practical significance for aided design, project selection of cable-stayed bridge and the coordinated effects evaluated between the bridges and surrounding environment when the bridges has been completed.

2. DESIGNING METHOD OF THE SYSTEM

2.1 The composition of the system

Based on WINDOWS operation platform and using MFC of VC++ and OpenGL, this paper develops the whole 3D visual simulation system of terrain and road. The system consists of four parts, the overall layout of the cable-stayed bridge, the cable-stayed bridge modeling, texture mapping and scene display respectively. The diagram of the system shows in Fig. 1.

2.2 Cable-stayed bridge with 3d modeling

There are three main techniques used for building 3D bridge models, including parametric input, artificial interactive input and graphics import. This system mainly uses the parametric input to establish the model of cable stayed bridge and has stronger interactive capabilities and higher efficiency for building modeling.

The basic functional units of the cable-stayed bridge include tower, cable, bridge, girder, pier, abutment etc. Its main design is through the key parameters to construct its modeling.

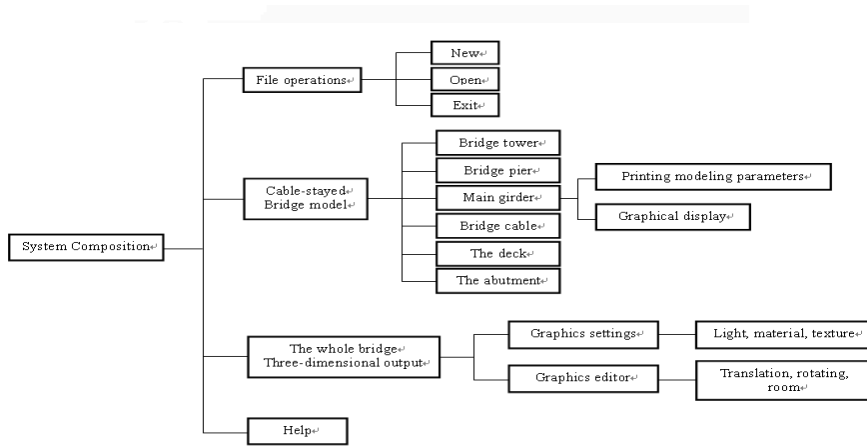


Fig1 The diagram of the system

2.2.1 Bridge tower modeling

Take the Tower parametric modeling as example to illustrate the modeling method.

Any 3D model can be abstracted as multi-layers including points, lines, surfaces and volumes. The points can be connected into line, the lines can determine the surface, the surfaces can be surrounded to volume, so the point is the most basic and the most critical level. OpenGL provides functions for painting point, line, surface and other basic graphic mapping. It's accomplished by calling `glBegin (GLenum mode)` and `glEnd ()` functions with different mode, the point, line and surface can be drawn respectively. The building process of tower modeling is mainly through the construction of tower polygons, polyhedral with a foot point of tower as reference point.

System composition Parameters in the program are as follows:

M_Tower_lenx (x is 1, 2, 3, etc.), // the length of the pillar section

M_Tower_widthx (1, 2, 3, etc.), // the width of the pillar section

M_Tower_hx (1, 2, 3, etc.), // the height of the tower, tower body, tower head respectively

M_Tower_cx (1, 2, 3, etc.), // the length of

spacing and beam between two tower pillars

$M_Tower_Beam_ex$ (x is 1, 2, 3, etc.), // the length of a beam section

$M_Tower_Beam_fx$ (1, 2, 3, etc.), // the width of a beam section

M_Tower_anglex (x is 1, 2), // angle of inclination of the control tower section

Fig2 shows the sketch map parameter of tower. In the design of the program, input tower parameters into the dialog, and the parameters can be changed to get different volumes of tower. Fig 3 shows such volume of towers obtained by running the program.

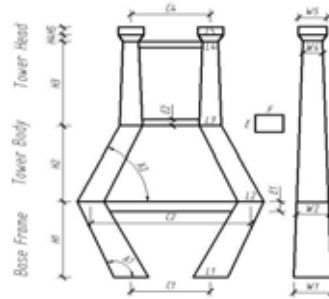


Fig2 Sketch map parameter of the tower

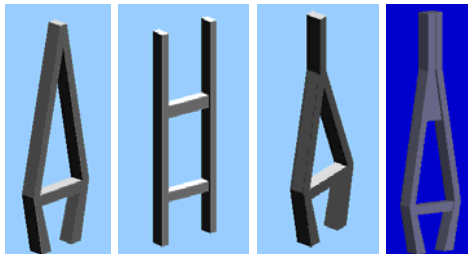


Fig3 Some familiar forms of tower

2.2.2 Girder model

Modeling process is similar to bridge tower, and this process is mainly composed of color fusion of OpenGL, the display list, and drawing a polygon, circling line drawing form the method and implementation.

The generated box girder and the plate girder are shown in Fig. 5, and 6.

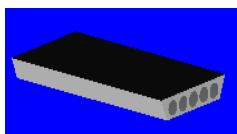


Fig4 Box girder

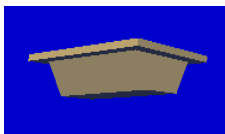


Fig5 Slab girder

3. AN APPLICATION CASE STUDY

Take the footbridge in Dalian Forest Zoo as example. The bridge is a concrete asymmetric cable-stayed bridge with single tower, the tower is a diamond type concrete tower, main span is double cable plane, anchor span is single cable plane, and the main tower adopts rectangular section, the main girder adopts solid sheet section. According to the above method to modeling bridge parts, input parameters and set light, then get the three-dimensional display image of the whole bridge which is shown in Fig 6. With the texture mapping technology of OpenGL, simulating cable-stayed bridge scene around and finally getting visualization result which is seen in Fig 7.

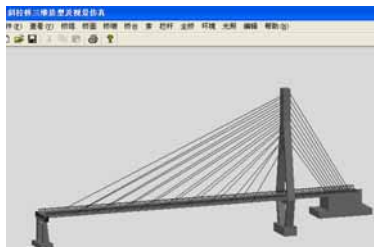


Fig6 Footbridge model in the zoo



Fig7 Footbridge with surrounding environment

4. CONCLUSION

This paper established 2D and 3D geometric parameters of the cable-stayed bridge 3D modeling system based on VC++ object-oriented programming, and use the core functions of OpenGL, practical library functions and computer graphics technics. Set the view, establish visual body, set light texture and carry the texture mapping in OpenGL, then carry out 3D visualization of the cable-stayed bridge. This study has an important significance in the early stage of design which includes bridge optimization and project selection.

REFERENCES

- [1] Tiejian LU, Youliang JIANG, Zhiwu YU, Research on three dimensional modeling and scene simulation. *Journal of Central South University of Technology (Science and Technology)*, 36(13):501-505, 6 2005.

- [2] Youliang JIANG. Research on and implementation of key technologies of bridge 3D-modeling and visual design. Central South University, School of Civil and Architectural Engineering, 2004.
- [3] Shiming XIANG. *OpenGL programming and instances*. Beijing, Electronic Industry Press, 1999.
- [4] Mingtao JIA , Linhui HOU, Changliang PAN, et al. Mining simulation system based on the OpenGL visualized components techniques. *Journal of Central South University of Technology*

through the postal address: School of Foreign Languages in Dalian Ocean University, Dalian, Liaoning, P.R. China, 116023.

ABOUT THE AUTHORS

1. Ziru WANG, is a Ph.D, professor, and the director of the Institute of Engineering Graphics in Faculty of Infrastructure Engineering Dalian University of Technology. Her research interests are Engineering Graphics , Application of Computer Graphics in the Hydro-power Resources and Civil Engineering. She can be reached by e-mail: wangziru26@126.com, by Fax:+86 (0411) 84674141,by phone : +86041184708525 or through the postal address: Faculty of Infrastructure Engineering Dalian University of Technology, Dalian, Liaoning, P.R.China, 116024.

2. Chunxia HU is a Master graduate student in Institute of Engineering Graphics in Faculty of Infrastructure Engineering Dalian University of Technology. Her research interests are Engineering Graphics. He can be reached by e-mail : liangyaolong@sina.com, by Fax: +86(0411) 84674141, by phone: +86041184708525 or through the postal address: Faculty of Infrastructure Engineering Dalian University of Technology, Dalian, Liaoning, P.R.China, 116024.

3. Mingqiu LI is a Associate professor in School of Foreign Languages of Dalian Ocean University. Her research interests are English Lexicology and English Teaching. She can be reached by e-mail: limq99@163.com, or

SUNSYS – A CASE STUDY ON GEOMETRIC COMPLEXITY IN COMPUTATIONAL DESIGN

Marco HEMMERLING and Jens BÖKE

Detmold School of Architecture and Interior Design,
Hochschule Ostwestfalen-Lippe, University of Applied Sciences

ABSTRACT: This paper presents a case study on geometric complexity that aims at an early integration of construction and fabrication parameters into the design process. Based on the academic project “SunSys”, the digital chain - from the initial form-finding to the final realization of spatial concepts - is discussed in relation to geometric principles. A process of transformation and optimization of the initial shape that integrated aesthetic, spatial and structural qualities as well as aspects of material properties and conditions of production followed the translation of geometric principles into a three-dimensional digital design model.

Keywords: Computational Geometry, Parametric Design, Architecture, Computational Construction, Digital Fabrication.

1. INTRODUCTION

The increasing influence and impact of digital technologies on architectural design and construction has become particularly visible in the past decade. These tools have allowed many architects for the conception and design of very complex architectural projects. Hence, formally intricate buildings have been designed and built over the past years. And many of these projects have undergone massive so-called post-rationalization-processes - i.e. methods of (mostly) intelligent geometrical simplification.

From today's point of view such processes, however extremely elaborate in themselves, appear to be a bit anachronistic. A more holistic potential of computers can be seen in the improvement of architectural projects in regard to their functionality and efficiency by considering and connecting important influencing parameters while negotiating requirements and constraints. Against this background the academic project SunSys focused on an all-embracing design approach, implementing and connecting all necessary information in a computational design process. Therefore, the early integration of optimization parameters, regarding structural performance, physical

properties and material specification as well as aspects of fabrication to inform the architectural design and construction were key aspects of the project.

2. DESIGN PROCESS

In order to match the overall requirements, various input parameters were developed and integrated into a parametric definition. They were established by analyzing onsite conditions using GPS tracking systems, image acquisition series and solar radiation analysis. In particular, existing patterns and structures of movement onsite, usage timeframes and alignment in relation to the course of the sun and to changing patterns of shade were used as source data for the design strategy.

The form-finding process developed generatively, while the values for the structures graduated succession of open and enclosed sections were derived from the findings of the time-usage analysis and the sun analysis. The profiles of the different usage areas, which form the pavilion's geometry, were created with reference to different seat positions, related to furniture – like a desk chair, bench, lounge- or deckchair.

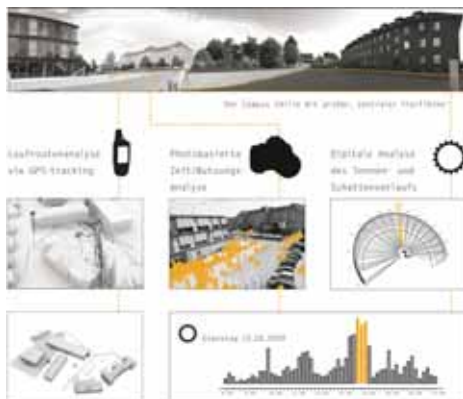


Figure 1: SunSys - analysis of influencing parameters for the form-finding process.

3. PRODUCTION PROCESS

After a selection process from multiple variations the final version of the pavilion was to be constructed from planar elements that are connected at an angle of 90° in order to simplify the assembly process. While all elements differ in shape, the CNC-fabrication of the building parts is based on a unified principle. The individual building parts, which also constitute the connective elements, are based on a parametric definition that incorporates all parameters, but also permits endless variations in terms of combination, allowing for an ideal form with a high degree of diversity and precision. Previously constructed surface objects define the overall shape of the pavilion, while the planar elements are programmed by a parametric definition in Rhino/Grasshopper3D.

3.1 Geometric constraints and definitions for production and assembly

The definition evaluates the initial surfaces and subdivides them by their UV-coordinates. In a following step, each coordinate is equipped with a plane, oriented to the normal-Vector of the surface. As an in-between result, a rectangular grid of perpendicular planes describes the pavilions geometry. To achieve the system logic, three different groups of planar elements

have been generated:

- The ones, which connect both other groups by being perpendicular on the guiding surfaces.
- A group of elements, connecting the perpendicular elements in the U-direction, by rotation around their x-axis.
- A group of elements doing the same for the V-direction by rotation around their y-axis.



Figure 2: SunSys - construction and assembly principle.

The rotation of the elements is defined by orienting every second plane, counting in both directions. The elements are generated as rectangular curves onto the planes. The perpendicular flat elements are created as guiding-elements; their dimension is for each direction half of the distance to the next evaluated point. It is important for the precision of the result to follow these guidelines, otherwise the planar interpretation of the double-curved geometry is leading to uncontrollable tolerances by just subdividing the distances.

In order to generate the rectangle curves of both other element-groups, a mid-point is generated on the rectangle-lines for each direction, being the start- or endpoint of a line, connecting to the midpoint of the next rectangle-line in that direction. By extruding the lines in the direction of their plane x- or y- direction, the rotated elements can be generated as rectangular surfaces. Accordingly, the distance of the extrusion is defining the geometry depth of the system.

For the connection of the elements, each

element has to be cut by half the distance of its intersection with the neighboring element (waffle-principle). Assembling both half cut elements will finally perform the full intersection of the elements. Through a Boolean operation the rectangles are extruded to solids.

3.2 Material aspects and tolerances

If the production process and the chosen material require the calculation of tolerances, two extrusions have to be performed: one without adding the tolerance, to define the final geometry element and one with an added tolerance to be just the cutting element, erased after performing the cut to its neighbor. In order to start a CNC-production process from planar panels the geometry of the resulting 3D-Model has to be unrolled.

In addition it is necessary to give identities to the 3D-objects that can also be milled into the physical pieces to keep control of the production and assembly process. In the first case study it appeared to be helpful to add information about the objects orientation, like a mark for the top face with an arrow indicating the assembling direction.

3.3 Generating variations

After this system is set up, variations can be applied easily. For instance by scaling the dimensions of the guiding elements following collected parameters, like the distance towards an attractor-point. Actually the form-definition of the SunSys-pavilion project depends on environmental parameters (solar radiation) following this logic.

4. CONCLUSIONS

SunSys should be understood as a prototype to achieve diversity and flexibility by computational efficiency. Given a different set of parameters, the design strategy would produce different overall shapes, functions and usage profiles, while the underlying geometric principle and fabrication method would remain the same. It is an illustrative example how to benefit from parametric design and the resulting

simplification of production and assembly processes without the need to post-rationalize geometric complexity.



Figure 3: SunSys - possible application of the system.

ACKNOWLEDGMENTS

The concept of the SunSys-project was developed by Jens Böke during his Bachelor thesis at the Detmold School of Architecture and Interior Design.

REFERENCES

- [1] H. Pottmann, A. Asperl, M. Hofer M., A. Kilian A: Architectural Geometry, Bentley Press, Exton PA, . 2005.
- [2] M. Hausschild, R. Karzel (ed.): Digitale Prozesse, Detail Verlag Munich, 2010.
- [3] M. Hemmerling, D. Lemberski: Sparkler – The Vitruvian Man vs. Buckminster Fuller in: Proceedings of the 30th eCAADe Conference, Technical University Prague, 2012.

ABOUT THE AUTHORS

1. Marco Hemmerling, professor of digital design and director of the post-graduate master's course on computational design and construction at the Detmold School of Architecture and Interior Design. Architecture diploma at the Bauhaus University Weimar. Post-graduate MA in architecture media management at the University of Applied Sciences Bochum. Member of the association of German archi-

tects (BDA) and the German association of craftsmen (DWB).

2. Jens Böke, PhD candidate at the Technical University in Delft. Coordinator of the post-graduate master's course on computational design and construction at the Detmold School of Architecture and Interior Design. BA at the Detmold School of Architecture and Interior Design. MA at the münster school of architecture (msa).

A SURVEY ON THE SPATIAL ABILITIES OF PRE-UNIVERSITY STUDENTS —ON SCORES OF MENTAL CUTTING TESTS—

Emiko TSUTSUMI¹, Toshikazu YAMAMOTO², Takeshi HONGO¹,
Hiroshi YANO¹ and Kenjiro SUZUKI³

¹ Otsuma Women's University, Japan ² Saitama University, Japan

³ National Institution for Academic Degrees and University Evaluation, Japan

ABSTRACT: We analyzed the average mental cutting test (MCT) scores of students in elementary, junior high, and senior high schools in order to investigate changes in their intuitive ability to understand three-dimensional shapes as a result of learning. The results are summarized as follows: (1) A significant difference was observed in the average MCT scores of students in different senior high schools. This difference was correlated with the deviation values for entrance examinations. (2) It was possible that the average MCT scores of the first grade students in junior high school was affected by the study for entrance exam in the previous year. (3) Comparison with the survey results obtained in 1992 indicated the possibility of decline in academic ability in the students in the present study. (4) Regarding the gender difference in the average MCT scores, a significant difference was found in the students in senior high schools. Although there was a gender difference in senior high school, the average MCT score of female students of the science course in a super science high school (SSH) greatly increased during the second grade, making it similar to that of male students. This significant increase might be explained by favorable influence of teachers' and female students' efforts.

Keywords: Spatial Abilities, Mental Cutting Test, Pre-university Students.

1. INTRODUCTION

Spatial ability is an important social ability. Potential ways to foster spatial ability through graphic science education in universities have been investigated since the 1980s. In particular, Suzuki reported the spatial ability of students as follows: (1) mental cutting test (MCT) scores greatly increased following lessons in descriptive geometry; (2) a high positive correlation was observed between the average MCT scores and the deviation values for university entrance examinations; and (3) the average MCT scores increased as graphic science education until senior high school progressed [1]. However, only a few studies have investigated the spatial ability of students in primary, junior

high, and senior high schools.

Regarding MCTs for pre-university students, Wu et al. analyzed students in primary, junior high, and senior high schools in Japan and China (approximately 50 male and 50 female students in each grade of primary, junior high, and senior high schools) and reported that, similar to university students, geometry education was important for pre-university students [2], [3]. Tsutsumi et al. conducted MCTs for female students in combined junior/senior high schools in 1995 and 2009 and analyzed the temporal change in MCT scores and the differences in the MCT scores between schools. The obtained results are summarized as follows [4]: (1) The results obtained by analyzing the differences in the average MCT scores between

schools based on the change in the governmental guidelines for teaching and the deviation values for entrance examinations suggested that the learning contents in the new governmental guidelines for teaching (revised in fiscal year (FY) 2008), in which the item of “projection” was revived in mathematics in the first grade of junior high school, might exert a favorable influence on spatial ability, which can be measured by relatively simple MCT problems.

(2) A high correlation between the average MCT scores of university students and the deviation values for university entrance examinations might exist for students entering junior high school.

(3) Lower accuracy rates for some problems in 2009 when compared to 1995 were observed in students who were in the second grade or higher of junior high school when the governmental guidelines for teaching (revised in FY 1998) were adopted, in which *Yutori* education (more relaxed education policy) was emphasized.

(4) In 1995, after the governmental guidelines for teaching (revised in the FY 1989) were adopted, in which basic education was considered important, the relationship between step-by-step learning and improvements in spatial ability was suggested. The accuracy rates of MCT increased with the age in line with the step-by-step learning. However, students receiving *Yutori* education had much slower increases (approximately one year) in accuracy rates for some problems than students receiving normal education. Moreover, the problems’ difficulty was not always in line with their accuracy rate.

Recently, the number of senior high schools adopting the accrediting system has increased, and the learning contents greatly differ according to school. Since the governmental guidelines for teaching have been revised about every 10 years, the learning contents also differ according to grade. The subject “mathematics” has been greatly affected by these revisions. We want to examine educational policies for acquiring abilities to freely manipulate spaces

and shapes, which are related to adaptation abilities in future occupations. For this, we should widely investigate the spatial ability of pre-university students. In other words, we should investigate the temporal changes in their intuitive ability to understand three dimensional shapes as a result of learning.

Therefore, we began to analyze the MCT scores of students in elementary, junior high, and senior high schools in 2011. The aim of the present study was to elucidate the changes in students’ MCT scores according to school and grade.

2. METHODS

2.1 MCT

In the MCT, a solid body and a contour line of a cutting plane were expressed using perspective drawings, as shown in Figure 1, and the true shape of a section was selected from five options. The MCT consisted of 25 problems (the full mark was 25 points). The MCT is a part of the Special Aptitude Test in Spatial Relation developed for university entrance examinations by the College Entrance Examination Board (CEEb) in 1939 [5].



Figure 1: Example of MCT problems.

2.2 Subjects

The study subjects were 423 fourth–sixth grade students in two public elementary schools (E), 1004 first–third grade students in four junior high schools (J), and 1533 first–third grade students in six senior high schools (S) (Table 1). Types and the name of schools and the number of subjects are shown in Table 2. In some schools, the test could not be conducted in some grades. Only Private_I was a combined junior/senior high school. Technical_G was a technical high school.

Table 1: Number of subjects in each grade.

Grade	E4	E5	E6	J1	J2	J3	S1	S2	S3
Male	59	79	65	208	241	36	401	358	165
Female	68	70	82	198	252	69	240	330	39

Table 2: Types and the name of schools and the number of subjects.

Type and name of schools	male	Female
Elementary school		
Public_W	99	112
Public_F	104	108
Junior high school		
Public_Y	140	141
Public_J	51	62
Private_S	216	164
Combined junior/senior high school		
Private_I	134	228
Senior high school		
Science course in public_O *1)	57	23
Humanities course in public_O *1)	198	164
Science course in girl's private_N	0	72
Science course in private_K	125	71
Integrated course in public_T	44	0
Humanities course in private_K	107	162
Technical_G *2)	337	41

*1) This school had been designated as a super science high school (SSH) between FYs 2005 and 2011.

*2) design, architecture, mechanical engineering, and electrical engineering courses

public.: public school

private.: private school

3. AVERAGE MCT SCORE

3.1 Grouping by difference in the average MCT scores between schools and courses

First, schools in which the average MCT scores have no significant differences were put together into a group.

(1) Elementary School

No significant difference was observed in the average MCT scores of students in the same grade in both sexes.

As the two public elementary schools, i.e., Public_W and Public_F, were located in the same district, no large difference existed in the educational policy between the two. Therefore, it was considered that the average MCT scores

of these two schools were similar to each other. Thus data for each grade in two elementary schools were put together into one group (code name: ES).

(2) Junior high schools

Regarding the male subjects, the average MCT score was significantly higher in Public_Y and Private_I than in schools Public_J and Private_S ($p < 0.01$). On the other hand, the average MCT scores showed no significant differences in Public_Y and Private_I and in Public_J and Private_S. Female subjects have the similar results.

The junior high school Public_Y is affiliated with a national university and has a high academic standard. In contrast Public_J is a municipal ordinary junior high school. Consequently, each of the male and female subjects were classified into two groups for junior high schools (code names: **HEIGHER** (Public_Y), **LOWER** (Public_J and Private_S)).

(3) Senior high school

A significant difference was mostly observed in the average MCT scores in different senior high schools, and even in courses in the same senior high school.

In Public_O, no significant difference was observed in the average MCT scores of the first-grade male and female subjects when comparing the science course and humanities course. However, in the second grade, the average MCT scores of the male and female subjects were higher in the science course than in the humanities course ($p < 0.05$).

Meanwhile, the average MCT scores were significantly higher in these two courses in Public_O than in other senior high schools ($p < 0.01$). Furthermore, there were some significant differences among these lower average MCT score schools.

In school Technical_G, no significant difference was observed in the average MCT score among four advanced courses (design, architecture, mechanical engineering, and electrical engineering courses).

With regard to senior high school, the devia-

tion value for the entrance examination was 73 for the science course in Public_O, 71 for the humanities course in the same school, 68 for the science course in Public_K, 63 for the humanities course in the same school, 63 for the science course in Private_N, 51–54 for Public_T, and 43 for Technical_G (additional note-1).

As described in Chapter 1, several findings (“a high positive correlation was observed between the average MCT scores and the deviation values for university entrance examinations, and the MCT score increased as graphic science education until senior high school progressed” (Suzuki [1]) and “a high correlation observed between the average MCT scores of university students and the deviation values for entrance examinations might exist in students when entering junior high school” (Tsutsumi et al. [4])) were already reported. Therefore, the deviation values for senior high school entrance examination might be correlated to the present average MCT scores among senior high schools.

Consequently, each of the male and female subjects were classified into five groups for senior high schools (code names: **SCI-1** (Science course in Public_O), **SCI-2** (Science courses in girl's Private_N and Private_K), **HUM-1** (Humanities course in Public_O), **HUM-2** (Humanities course in Private_K), and **TECH** (Technical_G)). Aside from this, combined junior/senior high school, Private I (code name: J-S), was analyzed as a separate group in order to examine the current of education consist of a set of processes through junior and senior high school.

3.2 Differences in the average MCT scores by grade

(1) Results

In Figure 2, the asterisk (*) added to the marker in the line graph expresses the results of significance tests between grades (one grade and a previous grade in a same group).

Regarding the male subjects, although the

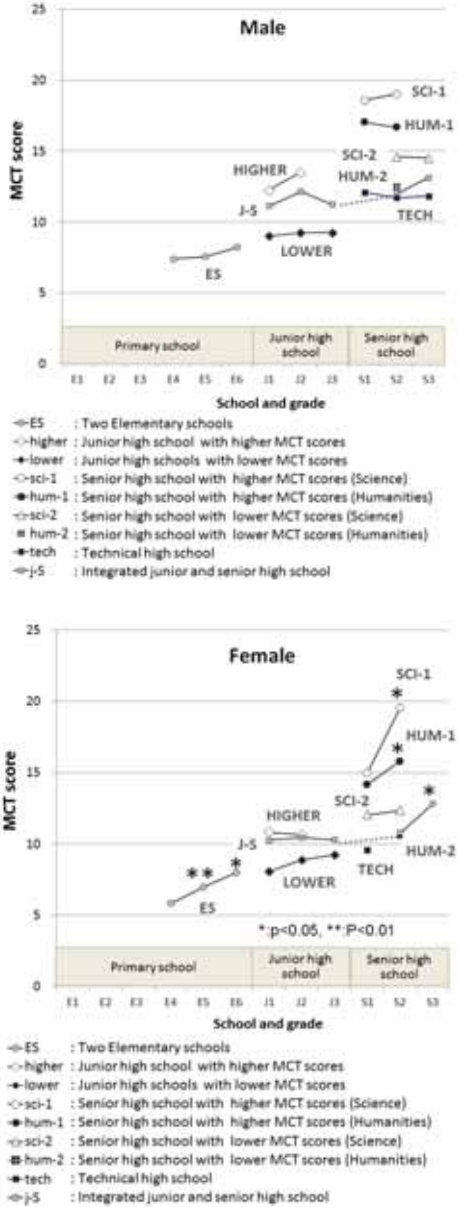


Figure 2: Average MCT score in each grade.

average MCT scores slightly increased as the subjects were promoted during primary and junior high schools, no significant difference was observed between grades. No significant increase was observed in the average MCT scores between grades in senior high school likewise (**SCI-1**, **HUM-1**, **SCI-2**, **TECH**).

Regarding the female subjects, their average MCT scores significantly increased between the fourth and fifth grades ($p<0.01$) and between the fifth and sixth grades ($p<0.05$) in **ES** (elementary school). Although the average MCT score slightly increased as the subjects were promoted during junior high school, no significant difference was observed between grades. The average MCT scores significantly increased in a period between the first and the second grades in **SCI-1** and in **HUM-1** ($p<0.05$), and also increased in a period between the second and third grades in **J-S**.

(2) Discussion

a. Differences in the average MCT scores of students in proceeding educational stages

In the present study, schools in proceeded educational stages were unrelated one another except **J-S**, which was a combined junior/senior high school as was described previously. However, the average MCT scores in proceeded educational stages were boldly compared with one another in order to understand the average MCT score trend.

When subjects in the first grade of junior high school was compared with the subjects in the sixth grade of **ES**, subjects' average MCT scores were significantly higher in **HIGHER** and **J-S** ($p<0.01$) both for male and female. However there was no significant difference in case of **LOWER**.

When subjects in the first grade of senior high school was compared with the subjects in the third grade of junior high schools, the average MCT scores were significantly higher only in **SCI-1** and **HUM-1** ($p<0.01$) for male and in **SCI-1** and **HUM-1** ($p<0.01$) and in **SCI-2** ($p<0.05$) for female. In case of the combined junior/senior high school, **J-S**, no significant

difference was observed between the average MCT scores in the second grade of senior high school and the third grade of junior high school.

The reason for this can be attributed to various factors, such as the existence of studying for entrance examinations and the contents of mathematics in the first grade of subsequent high school.

As for the effect of study for an entrance exam, **HEIGHER** with significantly-high average MCT score is a prestigious junior high school in the region and elementary students have to cram for an entrance exam. **SCI-1** and **HUM-1** with significantly-high average MCT scores have similar circumstances in entrance exams. In case of **J-S** which is also a prestigious junior high school in the region, subjects' average MCT scores were significantly higher at the first grade of junior high school. However, no significant difference was observed in the average MCT score in the second grade of senior high school and the third grade of junior high school. There is no exam for the students

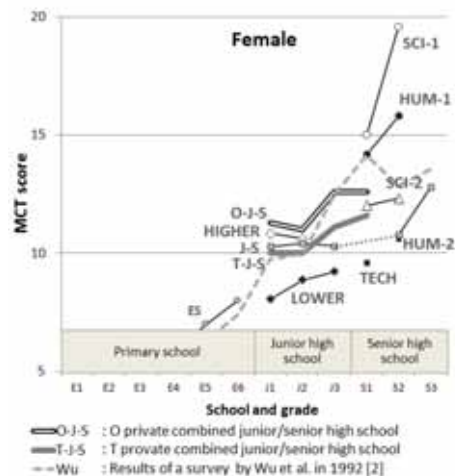


Figure 3: Comparison with the results in 2009

in the third grade to advance to its own senior high school. Thus it is possible that the average MCT score of the first grade students in the junior high school was affected by the study for entrance exam in the previous year. Or else, as was already presumed in the former study [4], these variations in the average MCT scores in the first grade students may be due to the existence of correlation between the deviation values for junior high school entrance exam and the average MCT scores.

The average MCT scores of students in two girls' combined junior/senior high schools in a period between the first grade of junior high school and the first grade of senior high school [4] were added and were shown in Figure 3 for reference using the double (O-J-S) and thick (T-J-S) lines without markers. In Figure 3, no large increases were observed in the average MCT scores in the period between the third grade of junior high school and the first grade of senior high school as was the case in J-S. The tendency was considered to be a characteristic of the students in combined school in which there are no exam for senior high school.

b. Comparison with the survey results obtained in 1992

In order to objectively examine the results obtained in the present study, the results of a survey conducted by Wu et al. in 1992 [2] were expressed by the dotted line in Figure 4.

Regarding the male subjects, the average MCT scores in the previous survey agreed with those of the present study for elementary schools. As for the junior high school, average scores in the previous survey was about the same as that of **HIGHER** and was significantly higher than that of **LOWER**.

Regarding the female subjects, although the difference in the average MCT scores between the present study and the previous survey was smaller than for the male subjects, the results appeared to be virtually the same.

The deviation value for the entrance examination for the junior high school was 56 for **HIGHER**, 60 for **J-S** and 54 for **LOWER** in the

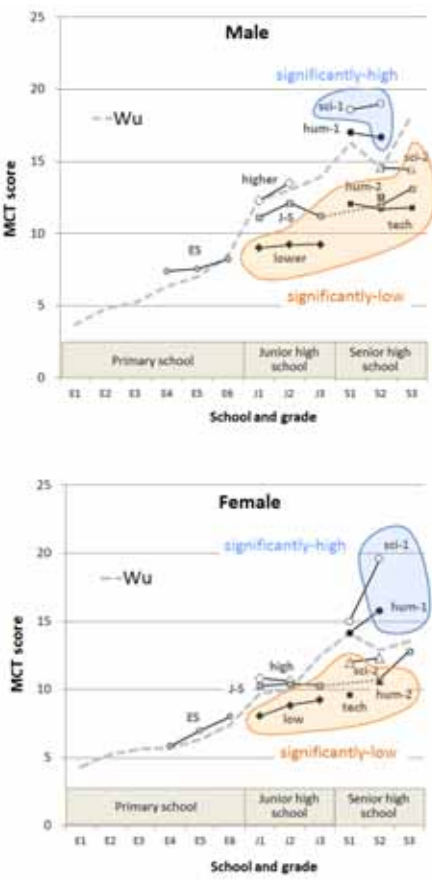


Figure 4: Comparison with the results in 1992 (Wu et al.) .

present study. In the previous survey, the deviation value for the entrance examination for the junior high school was only 50 or somewhat higher. Although the deviation values of the entrance examination were lower in the previous survey than **HIGHER** in the present study, average MCT scores of both surveys were about the same. The fact may indicate the possibility of decline in academic ability in the subjects in the present study.

With regard to the effects of the govern-

mental guidelines for teaching on the difference in the average MCT scores of pre-university students and between university students in the same grade have been pointed out. For example, Sugai et al. [6] reported that

“It is possible that education under the current guidelines increased the gap in scholastic ability between schools up to the high-school level, and scholastic ability in mathematics including graphic science related subjects fell even lower among students of Nihon University with its medium scholastic ability level than among students of the University of Tokyo with its high scholastic ability level, impacting the decline of spatial ability measured by the MCT.” Tsutsumi et al. [4] reported that “the learning contents in the new governmental guidelines for teaching revised in FY 2008, in which the item of “projection” was revived in mathematics in the first grade of junior high school, might exert a favorable influence on spatial ability, which can be measured by relatively simple MCT problems.”

The junior high school students in the previous survey and present study were both educated during the transition period of the governmental guidelines for teaching, which were announced to be revised in FY 1989 and in FY 2008, respectively. In both governmental guidelines for teaching, thorough teaching of basic educational contents and thorough acquisition of basic knowledge and skills were emphasized as keywords and have similar governmental policies. These two guidelines were being updated from the previous guidelines, in which *Yutori* education (relaxed education policy) was emphasized as a keyword of each. In the present study, junior high school students were educated under more relaxed education policy and educated with highly selected education contents when they were elementary school students. So they were strongly influenced by a *Yutori* education. This situation may also support the possibility of decline in academic ability in the subjects in the present study.

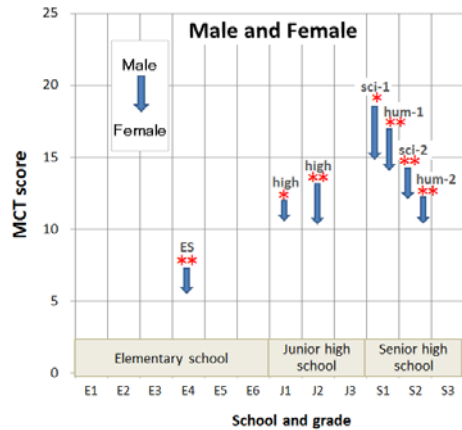


Figure 5: Gender differences in the average MCT scores.

In the senior high school, the average MCT scores extremely varied in the present study. Both in male and female subjects, the average MCT scores were significantly higher in **SCI-1** and **HUM-1** compared with the previous survey. **SCI-1** and **HUM-1** are the courses in a prestigious school in the region and junior high school students have to cram for an entrance exam.

3.3 Gender differences in the average MCT scores

(1) Results

Figure 5 shows the average MCT scores of the male and female subjects and the significance test results. The base of arrow and arrowhead expresses the average MCT scores of the male subjects and female subjects respectively.

During elementary school, male subjects' average MCT scores were only significantly higher than the female subjects' in the fourth grade ($p < 0.01$). During junior high school, male subjects' average MCT scores were significantly higher than female subjects' in the first and second grade only in group **HIGHER**. Regarding the senior high schools, although male subjects' average MCT scores were significantly higher than female subjects' in the

first grade in **SCI-1** and **HUM-1** by 3.6 and 3.5 points, respectively, no significant difference was observed in the second grade. In the second grade male subjects' average MCT scores were significantly higher than female subjects' in **SCI-2** and **HUM-2**, by 2.3 and 1.9 points, respectively.

(2) Discussion

a. Elementary school

Although female subjects' average MCT scores was significantly lower than male subjects' in the fourth grade of elementary school, their average MCT scores reached those of male subjects by sixth grade.

b. Junior high school

As shown in Figure 5, male subjects' average MCT scores were higher than female subjects' only in **HIGHR** during junior high school. Wu et al. [2] reported that the difference between male and female students' average MCT scores began to appear in the period between the sixth grade of elementary school and the first grade of junior high school. In the present study differences between male and female average MCT scores were observed only in school which has higher MCT scores.

c. Senior high school

Male students' average MCT scores were clearly higher than female students' during senior high school. However, the average scores of the female subjects in **SCI-1** and **HUM-1** particularly increased during the second grade and were almost the same as that of the male subjects in the same grade (Figure 2). In the data reported by Wu et al. [2], the average MCT scores of male and female students decreased during the second grade.

Figure 6 shows the relative frequencies of the MCT score of the first and second grade subjects in **SCI-1** and **HUM-1**. Regarding the subjects in **SCI-1**, although relative frequencies in the first grade varied widely, those in the second grade clustered on the higher side. In particular, the standard deviation of the female subjects in the second grade was as small as 3.26. Similar tendencies were observed in the

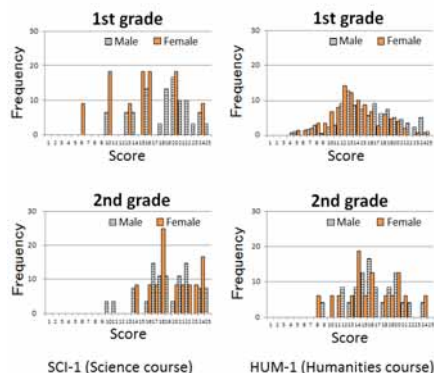


Figure 6: Relative frequency of MCT scores in **SCI-1** and **HUM-1**.

HUM-1. However, the tendencies observed in the **HUM-1** were not as clear as those observed in **SCI-1**. These tendencies might be due to the substantial educational contents in the science course, teachers' efforts to deepen students' interests in the field of science and mathematics, and students' strong desire to learn science and mathematics.

Regarding the **SCI-1**, one student noted that "preparation and review of the lessons are not easy, there are many tasks to be solved, and the club activities are very hard, but I spend a substantial life every day..." [7]. The results obtained in a questionnaire survey with students in **SCI-1** revealed that as many as 79% of the students wanted to be professionals, such as researchers in the field of science and mathematics and physicians, in which their personalities could be utilized. The students in **SCI-1** constructed a network related to science and mathematics in a prefecture their school was located and held lecture meetings to improve the desire for learning and the spirit of inquiry in addition to the school lessons. In other words, both male and female students planned various projects to deepen their interests in the fields of science and mathematics.

4. CONCLUSION

In order to investigate the temporal changes in the intuitive ability to understand three-dimensional shapes as a result of learning and to examine educational policies for acquiring the ability to freely manipulate spaces and shapes, which are related to adaptation abilities in future occupations, we began to analyze the MCT scores of students in elementary, junior high, and senior high schools in 2011. The present study revealed changes in MCT scores according to school and grade. The study subjects were 1612 and 1348 male and female students (a total of 2960 students), respectively, between the fourth grade of elementary school and the third grade of senior high school. The results obtained in the present study can be summarized as follows:

- (1) Based on the significance test results of the average MCT score at each grade in each school, classes in the same grade, whose survey results were similar to each other, were collectively analyzed.
- (2) Significant differences were observed in the average MCT scores of students in different senior high schools. This was considered to be due to the correlation between the deviation values for senior high school entrance examinations and the average MCT scores.
- (3) A slightly large increase of the average MCT scores was observed in the female subjects during the upper grades of elementary school. A large increase was observed at proceeded educational stages. It is considered that the average MCT score of the first grade students in junior high school may be affected by the study for entrance exam in the previous year. Or else, these variations in the average MCT scores in the first grade students may be due to the existence of correlation between the deviation values for junior high school entrance exam and the average MCT scores.
- (4) As for the gender differences, significant difference were found between the average MCT scores of male and female students in most senior high schools. However, in the sci-

ence course of a senior high school, which had been designated as an SSH, the female subjects' average MCT scores largely increased during the second grade, making them similar to the male subjects' scores. This might be due to the substantial educational contents in the science course, teachers' efforts to deepen the students' interests in the fields of science and mathematics, and students' strong desire to learn science and mathematics. Consequently, the female subjects' average MCT scores also greatly increased during the second grade.

Although further research still needs to be conducted, in the present study, the spatial abilities of pre-university students could be elucidated to some extent based on the MCT scores. In the future, we will investigate more students in combined junior/senior high schools in relation to governmental guidelines for teaching.

ACKNOWLEDGMENTS

The authors are grateful to the students and teachers for their participation in the present study during school hours. This paper is a summary of the results obtained in A Survey on the Spatial Abilities of Pre-University Students (the project number of 055) in the joint research project of the Institute of Human Culture Studies, Otsuma Women's University in FYs 2011 and 2012.

ADDITIONAL NOTE-1

The deviation value for the high school entrance examination (2011) used in the present study were survey data conducted by several entrance exam support services industry such as Koko Juken Navi, Chugaku Juken Navi, Yotsuya Otsuka and Shutoken Mogi Center.

ADDITIONAL NOTE-2

Certain upper secondary education schools are designated by the Ministry of Education, Culture, Sports, Science and Technology (MEXT) as "Super Science High Schools (SSHs)." Each SSH has developed enriched curricula, teaching

methods, and materials for science and mathematics in cooperation with universities and research institutes.

REFERENCES

- [1] K. Suzuki. Introduction to Cognitive Approach in Graphic Science (2): Evaluation of Students' Spatial Ability by the Use of a Mental Cutting Test. *Journal of Graphic Science of Japan*, 85, 5–12, 1999.
- [2] H. Wu and K. Suzuki. Development of Spatial Ability of Students during Elementary-, Junior High- and Senior High-School. *Proc. 15th International Conference on Geometry and Graphics*. 34:815-823. Montreal, Canada, 2012.
- [3] H. Wu and K. Suzuki. Evaluation of Development Process of Spatial Ability using a MCT (2): Survey Results of Elementary, Junior High, and Senior High Schools in Japan. *Proceedings of the 1995 Annual Meeting (Kansai) of the Japan Society for Graphic Science*. pp.101–106, 1995.
- [4] E. Tsutsumi and K. Watanabe. A Survey on Spatial Abilities of Female Students in Junior and Senior High Schools: Comparison with MCT Survey Results Obtained 14 Years Ago. *Proceedings of the 2011 Autumn Meeting of the Japan Society for Graphic Science*. pp. 119–124, 2011.
- [5] CEEB Special Aptitude Test in Spatial Relations, Developed by the College Entrance Examination Board, U.S.A, 1939.
- [6] Y. Sugai and K. Suzuki. Change over Time in Spatial Ability of Students Entering University – Impact of the Revision of the National Curriculum Guidelines up to High Schools–. *Journal for Geometry and Graphics*, 15(1):101–112, 2011.
- [7] Communication in Science and Mathematics,
http://www.ohmiya-h.spec.ed.jp/?action=common_download_main&upload_id=7405
 (obtained on November 23, 2012)

ABOUT THE AUTHORS

1. Emiko Tsutsumi, Dr. Eng., is a professor at the School of Social Information Studies at Otsuma Women's University. She can be reached by e-mail: tsutsumi@otsuma.ac.jp, or at the postal address: School of Social Information Studies, Otsuma Women's University, 2-7-1, Karakida, Tama, Tokyo 206-8540, Japan.
2. Toshikazu Yamamoto, Dr. Edu., is a professor at the Faculty of Education Saitama University. He can be reached by e-mail: tyamamot@mail.saitama-u.ac.jp.
3. Takeshi Hongo, Ph.D., is a professor at the School of Social Information Studies at Otsuma Women's University. He can be reached by e-mail: t_hongo@otsuma.ac.jp.
4. YANO Hiroshi is an associate professor at the Child Study Course, Otsuma Women's University. He can be reached by e-mail: yano@otsuma.ac.jp.
5. Kenjiro SUZUKI, Dr. Eng., is a Professor Emeritus of the University of Tokyo, and a Professor of National Institution for Academic Degrees and University Evaluation. He can be reached by e-mail: suzuki-k@niad.ac.jp.

THEORY AND PRACTICE IN THE IMPLEMENTATION OF AN ILLUSIONISTIC CEILING PAINTING AT PALAZZO MORONI IN BERGAMO, ITALY

Giampiero MELE¹, Giorgio BURATTI², Federica ROVO²

¹The University of Studies e-Campus, Italy ²The Polytechnic of Milan, Italy

*Machine eccelse, e sollevate Moli,
del Palaggio Moron son le grandezze;
eminenti cotanto ha le sue altezze,
che s'erge, al certo, à confinar co' i Poli.*

*Adorno di stupor, par che travoli,
A gareggiar del ciel con le vaghezze;
Le luci, a non salir, tant'oltre avezze,
riverenti al suo piè fermano i voli.....*

(Andrea Balioni Acad. Eccit. MI in D. CALVI, Le misteriose
pitture di Palazzo Moroni, Bergamo, Per Marc'Antonio Rossi,
1655)

ABSTRACT: The construction of Palazzo Moroni was commissioned by Francesco in the occasion of his marriage with Lucrezia Roncalli in 1631. The work lasted thirty years, from 1636 to 1666, and was executed by the master Battista della Giovanna. In 1649, Moroni entrusted to the painter Giacomo Barbelli from Crema (1604-1656) the decoration of some rooms. This is evidenced in the inscription written in the corners of the Gerusalemme Liberata hall: IO.s JACO.s BARBERI.s CREMEN.s INVEN.r PINGEB.t ANNO MDCLII. On the first floor of the building, you will find some amazing illusionistic artifices that impress with their high scenic quality. The illusions show architectural scenes that are able to expand the real space and to open trompe l'oeil ceilings using decorative elements. These examples of excellent workmanship are so specific that deserve to be studied in a national PRIN research on the perspective issues.

So, what is the hidden system used by these painters called "quadraturisti"? This work reveals, between hypothesis and validations, the operating stratagems to acquire and describe these works and the perspective reasoning used by the artist to give the illusion of having the sky in a room. Varied technical equipment has been tested to perform the survey. The 3D laser distancemeter measurements have been "dressed" with digital photographs edited with a software that calculates the intersection of optical lines from each photo and projected in three-dimensional space the generated points. The collected data allowed reconstructing the architectural structures with accuracy. The analysis of the digital model has revealed that the unit of measure used was the Bergamasco Foot (0.437767 m). The Gerusalemme Liberata hall measures 32x17 feet. The Golden Age hall measures 16x16 feet. The stairwell measures 34x12,5 feet.

The investigation on the ceiling perspective identifies in the first room four vanishing points instead that one. The comparison with the figure 100 in Andrea Pozzo's Treaty about the "modo di far la graticola nelle volte" (way to do the grids for painted vaults), has led to the hypothesis that Barbelli uses the same method. Pozzo explains that for flat artworks two grids are enough: one on the drawing, the other in the entire work. However, Pozzo himself criticizes the multiple vanishing points solution in a response notice to a raised objection. He states that "not everyone agrees that in a great perspective you give just one point to the whole work" and he replies that to fully enjoy and have an

optimal view of the whole work, it is more advisable to have a single point of view. Otherwise, you should look at the drawing from different points in the room, to perceive the whole work.

The research's goal is the geometric analysis and the restitution of the represented architecture. The purpose is to understand the practice for the construction of these kinds of architectonical paintings called "quadrature" and to highlight the peculiarities of the school established in Ducato di Milano. Padre Pozzo himself studied here, before he left the city in order to definitively settle in Rome (1681). The research analyzes a series of similar perspectives, with the aim to identify the elements that explain the theoretical training of Pozzo and his master.

Keywords: perspective, quadraturismo, painting.



Figure 1: The perspective ceilings of Palazzo Moroni in Bergamo

1. INTRODUCTION

In the upper town of Bergamo, under the Rocca Civica, stands Palazzo Moroni, authentic revelation of Baroque art in Lombardy.

Its construction was commissioned by Francesco Moroni in the occasion of his marriage with Lucrezia Roncalli in 1631. The work, which began in 1636 and lasted thirty years, were executed by the impresario Battista della Giovannaⁱ which is the performer of the work, while the designer's name was never known.

The four centuries of Lombard painting, represented in the ceilings of the rooms on the main floor of the building, include masterpieces by artists such as Giovan Battista Moroni, Bernardino Luini, Gian Giacomo Barbelli, Giuseppe Roncelli and Cesare Tallone.

The names of the Halls echo those of the subjects mostly represented in the frescoed ceilings during the Seventeenth century: the Golden Age, the Fall of the Giants and Hercules on the Carriage. In the dining room, even

ballroom since the Nineteenth Century, the myths beloved to the painters of the previous century gave way to the ventures of the heroes of the Gerusalemme Liberata by Torquato Tasso. In the rooms of the main floor you will find some amazing illusionistic artifices that impress with their high scenic quality. The “openings up” show architectural scenes that are able to expand the real space and to open trompe l'oeil ceilings using decorative elements, building an ornamental perspective that creates an imagined world. These examples of excellent workmanship are so specific that deserve to be studied in a national PRIN research 2010 “Architectural Perspectives: digital preservation, content access and analytics”. Bergamo’s case is useful to study the comparable ones in the Milan hinterland.

Among the paintings on the main floor of the palace, the Gerusalemme Liberata Hall excel. The name of the maker and the date of execution of this perspective representation, that transforms the real space, are documented by the entries listed in the corners of the room. The epigraph says “IO.s JACO.s BARBERI.s CREMEN.s INVEN.r PINGEB.t ANNO MDCLII”. This system of signing and dating is frequently used in paintings on ceilings or vaults. Here is revealed the name of the artist who depicted this view in 1652: Giacomo Barbelli was born in Offanengo (Cremona) in 1604 and died in Calcinato (Brescia), in 1656. He began his artistic career in the workshop of Tommaso Pombioliⁱⁱ. Barbelli, between the years Thirties and Fifties worked in Brescia, Crema, Lodi, Bergamo, and in the surrounding

villages. In 1649 Francesco Moroni relied on the decoration of the room of the Gerusalemme Liberata to the Crema’s painter. The iconography of the hall representation is described in detail by an Augustinian father three years after its completing. This is one of the few cases where it is possible to compare the painting with the narration of its representation.

2. THE ICONOGRAPHIC PROGRAM

The iconographic program of Palazzo Moroni is minutely described in the book “*Le Misteriose pitture del Palazzo Moroni spiegate dall’ansioso accademico Donato Calvi Vice Principe dell’Accademia degli Eccitati dedicato all’illustrissimo signor Francesco Moroni*”, published by Marc’Antonio Rossi in Bergamo in 1655. Donato Calvi, in the prologue dedicated to the Courtly Reader, declares that he was the inventor of the stories of the paintings. He wrote “...m’obligai alla spiegatione, onde da questa antecedente potrai dedurne fermissima la conseguenza, che non ambitione di lode, non pretentione di applausi, non intentione d’incontrar il genio di chi chi sij, habbi questi abozzi d’inerudita penna consegnato alle stampe, ma il solo fine d’adempir le parti d’un’ uomo di parola, & il solo stimolo di spiegar quel debito a che mi spingeva l’obbligo volontariamente contratto della promessa...”ⁱⁱⁱ. So, he had drafted that text as explanation of the paintings not for ambition of praise, applause claim or intention of appeasement. He printed these sketches of a not erudite pen, just to fulfil a personal commitment made with the customer.



Figure 2: Epigraphs above depicts name of the painter and the date of execution of the perspective of the hall “Gerusalemme Liberata”

In the major room of the palace, on the main floor, you can admire the painting that Calvi describes in his book and being impressed “*dall’invenzione d’una maravigliosa prospettiva di sotto in su superbamente arricchito, sembra inalzarli verso le stelle, perche verso le stelle ogni mente per l’ammirazione si sollevi*” (the invention of a wonderful rich perspective from the above upward: it seems to raise the characters to the stars, as each mind has to elevate till sky for admiration)^{iv}. The painting, which depicts emblems, hieroglyphics, statues, ligatures, enterprises and trophies, is meant to describe the deeds of Gerusalemme and at the same time the glorious story of the Moroni’s family, reported in the four corners of the hall.

The Gerusalemme Liberata is the most important epic poem by Torquato Tasso and tells the fighting between Christians and Muslims at the end of the First Crusade, during the siege of Jerusalem. The work, published in 1575, revolves around Goffredo di Buglione, the leader of the First Crusade, that after six years of struggle awaits the end of winter, when he receives the apparition of the Archangel Gabriel

that helps him to take command of the army and lead the final attack against Jerusalem, at that time occupied by the Muslims. Indeed, the painter Gian Giacomo Barbelli began to paint in the three central spaces of the ceiling the mission of the Archangel helping Goffredo, in accordance with can be read in the first canto.

Around the room, at the base of the ceiling “*...spicca con l’arte del pennello in disposizione dorica un bellissimo Cornicione*” run the painted frieze of a beautiful Doric cornice which raises “*all’in su un’eminente loggia da ordine duplicato di colonne compartita...*”, a prominent lodge consisting in a double order of columns^v. The coupled columns rest on a bronzed and golden balustrade and support the vaults which frame the representations of some of the Goffredo’s undertakings described by Tasso. The facts are explained in eight emblems painted under the arches. In the eight Serliana of the lodge are represented the most famous actions that Tasso wrote in his poem: the review of the army of Christ (1); the compassionate case of Olindo and Sophronia (2); the tricks of the sagacious Armida (3); The de-

vout procession around Jerusalem (4); the combat and the unfortunate death of Clorinda (5); the value of Tancredi in the enchanted forest of Ismeno (6); the captivity of Rinaldo within the garden of Armida (7); the taking of Jerusalem and its fall in the hands of Godfrey (8).

Between a column and another are represented eight standing bronze statues, which are the eight virtues that have accompanied the great enterprise. Four are placed on the corners of the room, they wear female clothes and they are Faith (A) Skillfulness (D) Strain (E) and Victory (H). The four that are in the centerlines of the two long sides are the masculine virtues: Counsel (B), Zeal (C) Disdain (F) and Rejoicing (G). The perspective is decorated with figures of youths wielding war trophies (sword, golden or laurel or spines crowns). In the midst of these, twelve crowned heads are painted in small niches, showing the eight kings who suc-

ceeded Godfrey in the Kingdom of Jerusalem and other heroes who took part in the conquest of the Holy City.

In the four corners of the vault are lastly painted the accomplishments of Moroni's family. The first depicts the fortress inside which a blackberry tree is planted to indicate the marriage between Francesco and Lucretia Roncalli (α); in the second you find a blackberry plant in the sunlight, to signify that the line of Moroni away the darkness and shadow of imperfection (β); in the third painting there is a blackberry twig laden with fruits that represents the growth of Moroni family originated by honor, merit and work of the entire lineage and not by luck (γ); in the fourth corner is depicted a blackberry branch between a wooden wheel and the framework of a ship, these symbols tell anyone that Moroni protects like a ship against storms, if the luck helps (δ)^{vi}.



Figure 3: The eight actions described by Torquato Tasso in the Gerusalemme Liberata.

In the iconography the number is important, in fact it is remarkable that the author distributes emblems, statues, columns, according to

non-random numbers based on number four. You can see 4 coats in the corners, eight enterprises of the Gerusalemme Liberata, 8 statues

of the virtues, 12 busts of kings, 16 young boys and 32 columns that generate the loggia.

So, what is the hidden system used by this painter affiliated to the "quadraturisti", to place in that ceiling what has been described? To answer this question it has been necessary to ad hoc set up a survey that would allow to develop a model that could be helpful at this type of analysis.

3. THE INTEGRATED SURVEY OF THE CEILING OF THE GERUSALEMME LIBERATA HALL (GIORGIO BURATTI)

The survey model of the ceiling of Palazzo Moroni has been obtained by integrating more indirect methods of measurement. The method of three-dimensional photogrammetry was considered the most appropriate to the purposes and characteristics that the model had to comply. The two-dimensional photogrammetric straightening that would be obtained by some appropriate software, would be equally satisfying to produce a reliable metric model, however, we chose to use the three-dimensional photo-modeling to obtain a 3D model that would allow both to learn about the deformations of the ceiling in all its points, and to make known the work in a more realistic perception.

The 64 pictures used were taken with a Nikon D3100 camera equipped with AF-S Nikkor 18-55 mm lens. Setting the machine to "Auto" mode, the images have been taken by putting the body on the ground, with the only care to place the shooting points along three parallel tracks. This has permitted to have frames with a good overlap (more than 30%) and with an uniformity of parameters.

The software used to process the images, Agisoft PhotoScan Pro, gives back a dense cloud of 214.532 points, which generate a polygonal model of 23.308.247 triangular faces. The mesh has been geo-referenced and high-definition textured. The almost fully automated workflow permits to a non-specialist to handle any images and to have high-quality photogrammetric data. The speed and ease of

the detection operations combined with the quality of the 3D model mapped with photographic images with uniform colors, make this type of survey appropriate to the description of painted ceilings and vaults, with the purpose of studying their illusionistic perspectives.

The photo model was supported by a model taken by Leica 3D Disto laser-meter obtained by measuring a floor plan at the level of about 1 m, two sections in the longitudinal and transverse centerline, and more than 30 remarkable points of the painted architecture. The merge of these two methods has been completed in 3D virtual space of Rhinoceros 5.0, in which you can import both the texturized .obj format generated by PhotoScan, and the .dxf obtained from 3D Disto. The two models must be manually aligned.

The final model has high metric quality and it consents any type of integration and study. The analysis of this model leads to the identification of the hypothetical point of view of the perspective used by the painter to create that excellent optical illusion.

4. THE STUDY OF THE "QUADRATURE" BETWEEN ANALYSIS AND TREATISES (GIAMPIERO MELE)

Before getting to the heart of the analysis of the ceiling perspective in Gerusalemme Hall in Palazzo Moroni, it is useful to explain what is meant by the term "*quadratura*". What is called *quadraturismo* is still uncertain: in mural painting, the historiography itself does not operate a sharp distinction between illusory architectures and *quadratura*. However, you could define the first as a representation on a wall or a ceiling, through drawing, color and the laws of perspective, of a three-dimensional space, including architectural elements. Instead, the second is an artistic genre that combines elements of painting and the architecture, mainly ceilings or vaults, able to change the perception



Figure 4: The survey of the ceiling at Palazzo Moroni in Bergam.

fusion with the representation of an architectural space imagery conceived ad hoc.

According to this definition, in the case of the painted ceilings of Palazzo Moroni, can we call them murals or quadrature? To answer this question requires a good analysis of the depicted perspective, with the aim of identifying the method by which the painter has illustrated the architecture. Then, it is necessary to check if that ones have a proportional relationship with the real space that contains them.

The first step is to identify the vanishing point or points, scientifically hypothesize the location of the point of view of the perspective, to build backwards a restitution of the architecture obtaining elevation and plant sections. Finally, you can see if the proportions of the represented architecture match those of the real

one.

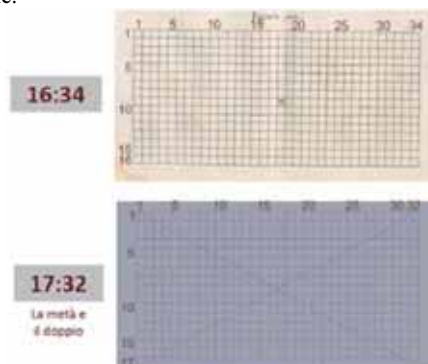


Figure 5: Image 1 represents the grid of fig. 100 contained in the Pozzo's treatise. Image 2 is the grid identified for the drawing of ceiling in Palazzo Moroni (Bergam).

The measurements in meters obtained from the survey, carried out by the authors, converted into Bergamo's Feet^{vii} express the numerical proportions of the sides of the parallelepiped identified by the room. The numbers associated with the sides of the hall are useful to the identification of those that Father Pozzo defines "grids"^{viii} and uses to explain the projective system of a perspective on ceilings or vaults. The grid in the figure 100 of his treatise has the proportion 16:34. The size of the room of Gerusalemme is 32 x 17 Bergamo's Feet. So, the sides of the grid used by Gian Giacomo Barbelli are double and half of those shown in Figure 100 by Pozzo mentioned above. The height of the room is $14 + \frac{1}{3} = 6.2746$ m. The hall of the Golden Age measure 16x16 Feet and is again $14 + \frac{1}{3}$ Feet high. The staircase measures 34x12.5 Feet and has a height at the landing on the main floor equal to that of the rooms.

A detailed analysis of the perspective from below upward of the ceiling of the Gerusalemme Hall, also by drawing the lines corresponding to the vertical edges of the architecture, leads to the first observation that, compared to Pozzo's treatise, the perspective does not have a single point of view. Four vanishing points are identifiable and their position is related to the vertices of a quadrilateral obtained by crossing two equilateral triangles.

In the description of Figure 100, the Jesuit includes the response to an objection that was made to him: "*Non tutti approvano che in una gran prospettiva si dia un sol punto a tutta l'opera, per esempio alla nave di mezzo del tempio espresso nella fig. 93 non consentono si assegni un sol punto ma vene vorrebbero molti*". He says that not everyone approves that in a great perspective you have a single vanishing point, but they would like to see many of them. Pozzo continues by saying "*se per esempio in una volta dove vogliate dipinger un solo corpo unito d'architettura, e figure voi ponete più punti di veduta, non avrete alcun luogo d'onde possiate goder tutta l'opera, ma vi conviene girarla per tutte le parti e goderla al più a poco*

a poco". If, for example, in a vault where you want to paint a single architecture and figures, you place multiple points of view, you'll have not a location from which to see all the work, but you need to turn around and enjoy it little by little.

Then, father Pozzo prefers the solution with a single point of view. His objector prefers more points of view, with the division of the space in a similar number of parts. In the stairs of Palazzo Moroni you find this solution. However, regardless of this, the question suggests that is common to build "opening up" perspectives with multiple points of view, not caring that these correspond to the projection point to have a perfect deceive from specific points.

The ceiling of the Gerusalemme Hall is similar to those criticized by Pozzo, but it is probably one of the most widespread cases in the construction of illusionistic perspective. One of the writers, though not Italian, which describes the solution seen at Palazzo Moroni is Antonio Palomino de Castro y Velasco (1653-1726)^{ix}. In his treatise "*Il Museo Pictorico y Escala Optica*" in two volumes, published in Madrid in 1723, in the second volume, in Book Eighth, in section V, he describes the fig. 4 of the Lamina 10. Exposed here is the "*practica par a mayor inteligencia de la perspectiva de techos*" and "*planta ma comprehensible de la perspectiva de techos*"^x. Here the author describes in detail the system to achieve the heights in the perspective from below "at the Italian manner". He does not explain the positioning of the four vanishing points, that are placed in the same geometric relationship of those in the ceiling of Palazzo Moroni. It explains in detail how to determine the height to characterize a closed-in courtyard architecture. In the description at p.181 of his treaty, Palomino shows the rectangle of the room in which to paint a planar perspective on the ceiling and to identify the main point in the center of the surface through their diagonals.

Considering a quarter of the underdrawings of perspective laying on the floor, the vanishing

points of the vertical straight lines in the picture will be located on the middle of the room. About the identification of the heights he says: *"Pretendo yo levantar un edificio sobre la linea d, h, che tenga de altura tanto como la misura d, h, para lo qual tiro desde el angulo d, la linea punteada à el punto de la vista e; y despues desde el punto h, tiro a el punto de la distancia, che està fuera, la linea h,K, que cortará la d, e, en j, y hasta allí serà el escorzo, ù degradacion justa de un edificio de la altura d,h,...y tirando la j,m, paralela à d,h, darà la degradacion de todo aquel costado, contenida en el trapecio d,j,m,h"*.

So nothing is said about the four vanishing points. However, we can know some information that we cannot find in Pozzo's treaty: the first grid (the preparatory drawing) must lay on the floor, and that to assign any height, you have to set that dimension on the long side of the room and draw from this point with the straight line converging to de point of distance K. Therefore, the system to identify the perspective on the plane ceiling is revealed merging the information from these two treatises, written twenty years apart from each other^{xi}.

At this point it is easy to find the height of the architecture painted in the Gerusalemme Hall. In fact, if you consider the point of intersection of the straight line parallel to the longer side of the room, which depicts the reduced side of the architecture represented, and intersect the diagonal, you get the point **j**. From this you draw the straight line **Kj** end extend it until you intersect the long side of the room at the point **h**. The connection point identifies the height assigned to the represented architecture. This number, in our case, corresponds exactly with the 2/3 of the long side of the room that measures 21+1/3 feet^{xii}.

Once seen the process to find the height, it remains to clarify the way Barbelli used to draw the preparatory design on the ceiling. About it, Pozzo writes that for planar paintings, two grids are enough, one in the drawing, the other in the

work. Palomino also declares that the preparatory drawing, a quarter of the room, should be placed on the floor. From the two texts you can deduce that the first grid, a prospective drawing, must lie on the ground. Then, still following Pozzo's indications, you foresee you have to attack a strong thread in the vertical center of the room from

In conclusion, in the "opening up" perspective of Palazzo Moroni, the point of view has not been properly calculated and used as a projection point to get the second grid, as Pozzo supposed. In this case, the position of the projection point is a function of the size of the underdrawings, which is proportional to the room's measures. It follows that if we consider the division between mural painting and quadratura made at the beginning of this paragraph, the Bergamo perspectives cannot be considered like that.



Figure 6: Views of the Palazzo Moroni hall.

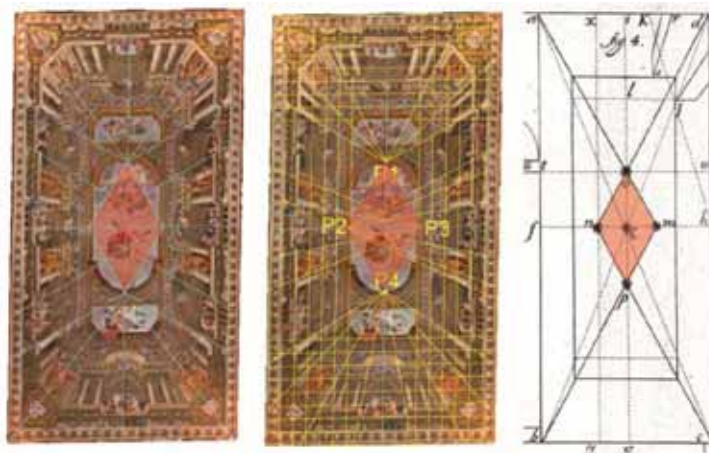


Figure 7: From left to right: Individuation of vanishing points with the double equilateral triangle. Grid obtained with the foot of Bergamo. Illustration contained in Palomino's treatise for the drawing of ceiling perspective

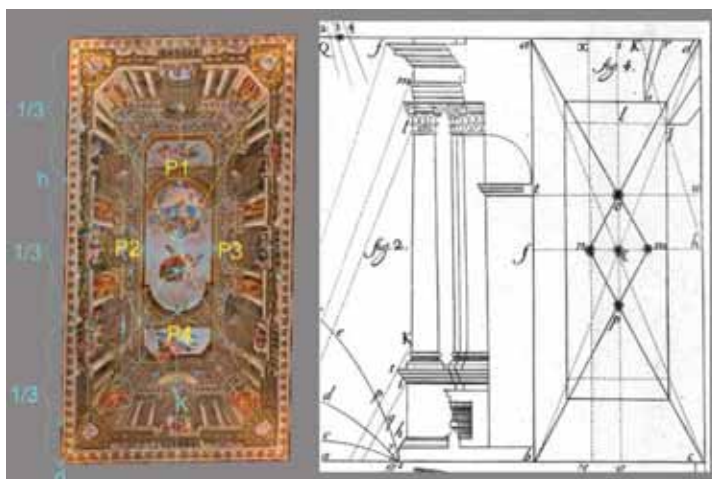


Figure 8: From left to right: the rule for the identification of heights in "Gerusalemme Liberata". Illustration contained in Palomino's treatise for the drawing of ceiling perspective

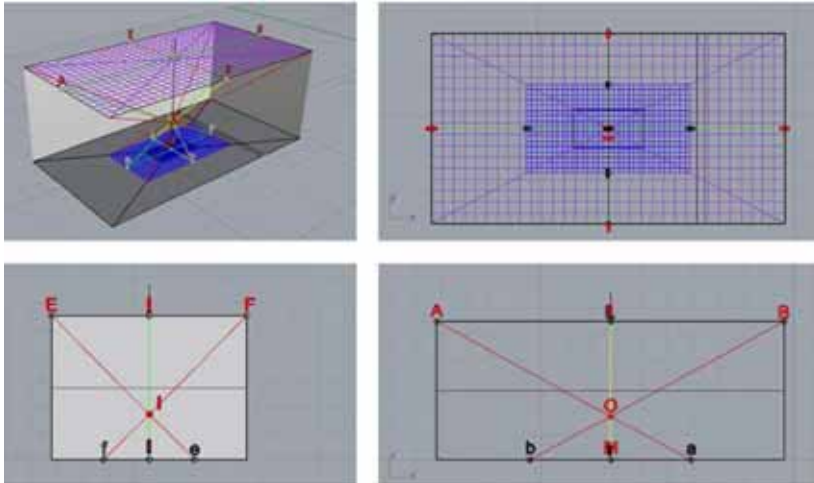


Figure 9: Pattern for projection of grid on ceiling paintings: one drawing is on the floor and the other in the ceiling.

5. REFERENCES

- [1] D. CALVI, “Le Misteriose pitture del Palazzo Moroni spiegate dall’ansioso accademico Donato Calvi Vice Principe dell’Accademia degli Eccitati dedicato all’illustrissimo signor Francesco Moroni”, *Per Marc’Antonio Rossi*, Bergamo, 1655.
- [2] G. COLOMBO, M. MARUBBI, A. MISCIOSCIA, (a cura di) “Gian Giacomo Barbelli: l’opera completa”, *Edizioni Grafica*, Ombriano, 2011.
- [3] A. MARTINELLI, (a cura di) “Andrea Pozzo”, *Electa*, Milano 1996.
- [4] A. PALOMINO DE CASTREO Y VELASCO, “El Museo pictórico y escala óptica”, en *La imprenta de Sancho*, Madrid, 1723.
- [5] A. POZZO, “Prospectiva pictorum et architectorum”, *Typis Joannis Jacobis Komarek Bohemi apud S. Angelum Custodem*, Roma nel 1693.

ABOUT THE AUTHORS

1. Giampiero Mele is architect and full Professor of drawing, received his Ph.D. in “Survey and Representation of architecture and the environment” at the University of Florence in 2004, and a Ph.D. in “Architectural and Urban Design” from the Université di Paris 8 in 2004. From 2010 to 2014 he was researcher at the University of Study eCampus, and professor of descriptive geometry at the University of Ferrara. His fields of research are the relationships between geometry and arithmetic in historical architecture, and drawing in architecture and design. He has given talks at various national and international conferences in these fields, and is the author of numerous papers, the most important of which are: “Mesure et proportion dans la Loge de la

- Signoria à Florence” (Revue XYZ, Association Française de Topographie, no. 98, 2004); G. Mele, A Geometrical Analysis of the Layout of Acaya, Italy. NEXUS NETWORK JOURNAL, vol. 14, p. 373-389, Birkhauser, Springer 2012
2. Giorgio Buratti is a PhD student at the DERE unit (Design and Representation) in Design School at the Polytechnic University of Milan. He graduated with top marks from the same university where he obtained a master's degree in ergonomics too. He has participated in numerous national and international conferences in the field of design and ergonomics and he has written several scientific publications. His research topics concern the ergonomic design and the influence that the new tools of digital drawing have on the design process, in particular he studies the parametric generative processes applied to complex morphological systems.
 3. Federica Rovo is junior architect and graduated in architecture at the Polytechnic University of Milan. She collaborates as assistant professor with prof. Giampiero Mele at University of Study eCampus and honorary fellow of drawing classes. Main interests are directed to the disciplines of representation and to survey analysis through historical measures.
-
- ⁱ As you can see in the “Libro dei conti della Fabbrica di Porta Penta” kept in the Moroni Family Archive.
- ⁱⁱ Cfr. G.COLOMBO, M.MARUBBI, A. MISCIOSCIA.
- ⁱⁱⁱ Cfr. D.CALVI, pag. 53.
- ^{iv} Cfr. ibidem, pag.53.
- ^v Cfr. ibidem, pag. 54.
- ^{vi} The Fortune is allegorically represented as a wagon wheel.
- ^{vii} The Bergamo’s Feet structure (43,7767 cm) is closed and composed by multiples (Braccio da fabbrica (53,1414 cm), Braccio Mercantile (65,9319 cm), Cavezzo = 6 Ft) and sub-multiples (Oncia = 1/12 Ft, Punto = 1/12 Oncia, Atomo = 1/12 Punto).
- ^{viii} In the second volume of Pozzo’s Treatise “*Perspectiva Pictorum et Architectorum*”, printed in 1700, in fig. 100 the system to do the three grids is described.
- ^{ix} Spanish learned treaty’s writer, architect, decorator and painter.
- ^x Cfr. A. Palomino, pag.181, 182
- ^{xi} 1703 Pozzo, 1723 Palomino.
- ^{xii} The major side of the room is equal to 32 Ft.

THIRTEEN YEARS OF MRT – RESULTS, OPTIONS AND DILEMMAS

Domen KUŠAR and Mateja VOLGEMUT

University of Ljubljana, Faculty of Architecture, Slovenia

ABSTRACT: During the 2013-2014 academic year, the 13th annual test of spatial ability is being performed at the Faculty of Architecture of the University of Ljubljana. The first goal of the test is to obtain information about spatial ability of first-year students. Thus, the test is performed at the beginning of the academic year and covers the population of first-year students. Several ways of checking spatial ability exist, since this is a complex field that contains various aspects of spatial ability. At the Faculty of Architecture, we use the Mental Rotation Test which is one of generally recognized tests for checking spatial ability. The conditions of the test have stayed the same throughout the years (time, exam papers etc.) which enables us to objectively compare different generations. Considering the structure of the test, its primary goal is to check the capability of mental rotation, which is an integral part of spatial ability. More than 2000 ($n=2094$) students have performed the test up to now. Besides basic data about generations, for the past three years we have also been comparing the differences in terms of left-handedness and right-handedness. The results show no statistically provable differences. But it is true that the tested population of left-handed students is, despite the above, relatively small ($n=66$) even though it proportionately does not deviate from the share of the entire population. An interesting phenomena is also the poorer results of each third generation, which has been appearing throughout the years. This difference has been statistically proven for 2009 and 2012, while the difference for 2003 and 2006 has not yet been statistically proven.

Processing of results with the help of statistical software makes it possible to check various hypotheses from the area of spatial ability and provides interesting conclusions, presented in the article. A large digital database also makes it possible to evaluate the test and the students' approach towards solving it. This opens up new possibilities in the area of comprehensively treating spatial ability and, at the same time poses new questions and dilemmas. One such dilemma is the existing scoring system. Results obtained in the joint project of the Faculty of Architecture of the University for Technology and Economics in Budapest, Ybl Miklos University for Architecture and Civil Engineering of the Szent Istvan University of Budapest and the Faculty of Architecture of the University of Ljubljana namely show that the existing scoring system is not the most appropriate since it does not provide the envisaged normal distribution of the sample. Based on obtained data, a new scoring system has been proposed which enables a normal sample distribution.

Keywords: MRT (Mental rotating test), spatial ability, Faculty of Architecture

1. INTRODUCTION

Human capacity comprises different types of intelligence: linguistic, mathematical, kinaesthetic, naturalistic, musical or personal and spatial. A component part of the latter is spatial

ability. The study of spatial ability is a relatively young discipline. A pioneer in this field was Sir Francis Galton [14], who reported about his experimental inquiries into mental imagery in 1880. Subsequent research was characterised by

the confirmation of the existence of spatial ability separately from general intelligence, which was achieved with the help of psychometric studies. This was followed by the confirmation of the existence of several component parts of spatial ability. This also caused some confusion due to differences in terminology as regards these component parts, which was probably due to the use of different techniques in the analyses. J.P. Guilford [3] thus established the following components of spatial ability:

1. spatial orientation, which comprises the ability to perceive three-dimensional space and objects shown on a two-dimensional medium,
2. the ability to present a three-dimensional object on a two-dimensional medium and the ability to determine the entire composition on the basis of knowing only fragments,
3. the ability to search for spatial solutions with the help of drawings and the ability to express oneself in three-dimensional space according to verbal instructions,
4. the ability of fast perceptions and actions in space.

McGee [13] combined these elements into two strong groups: spatial visualisation and spatial orientation. He also established that it was difficult to set a strict line of separation between individual components of spatial ability. However, by utilising their studies, Linn and Peterson [12] listed the following definitions of terms related to spatial ability:

1. spatial ability is a general skill for representing, transforming, generating and recalling symbolic, non-linguistic information. Spatial ability encompasses mental rotation, spatial perception and spatial visualisation,
2. mental rotation encompasses the ability to mentally rotate two or three dimensional figures or objects quickly and

accurately,

3. spatial perception is an individual's ability to determine spatial relationships with respect to the orientation of one's own body despite distracting information,
4. spatial visualisation is the ability to manipulate complex spatial information. All this requires the analysis of connections between different spatial depictions.

In addition to these four, there is supposedly also a fifth term, i.e. dynamic spatial reasoning, which includes judgements on relative speed and distance [10].

Based on analyses of spatial ability, Yilmaz [26] prepared a model that represents the main components of spatial ability. These are:

1. Closure speed,
2. Spatial orientation,
3. Environmental ability,
4. Flexibility of closure,
5. Spatial relations,
6. Spatial visualisations,
7. Spatiomemoral (spatiomeporal) ability,
8. Perceptual speed.

The term spatial ability thus covers a broader area that is comprised of a number of different skills. Consequently, we are dealing with a terminology that combines different combinations of the words visual and spatial with the words ability, skill, orientation, perception and similar, all with the aim of trying to define in as much detail as possible every element of spatial ability.

Unfortunately, spatial intelligence was underestimated compared to the other previously mentioned intelligence types [19], which was a false determination and this has been proven already by Piaget and Inhelder [16] and Huttenlocher and Newcombe [5]. This is also corroborated by newer studies [4]; [2].

The perception of space is in the domain of the right side of the brain [23], and it has been

known for years that the right hemisphere is bigger in men than in women and also develops sooner [9]. This is probably also the reason for the differences between the genders in results of spatial ability tests. The difference is most obvious in mental rotation tasks, less obvious in orientation and not at all in visualisation [12]. The majority of experts agree that differences are not noted until after puberty and that growing up has a major effect on the development of this ability [15].

The results and new findings in the field of spatial ability increase interest in the field. The development of the society, different conditions of growing up and above all the effects of technology on all areas of our lives probably also indicate changes in this field. It is also known that good spatial ability is indispensable in everyday life and absolutely necessary in specific jobs such as in the fields of architecture, construction, mechanical engineering, the military, specific sports, transport and others.

Special tests for assessing spatial ability have emerged from the desire to improve this ability and methods to improve it have been created. Developed economies have namely come to realise the enhanced financial effect that followed improved spatial abilities of employees. The methods towards improving it are close to the methods used by traditional descriptive geometry for dealing with spatial relationships and for projecting elements of space onto different media. More extensive studies have established that courses in descriptive geometry have a favourable effect on spatial ability. This has been proven with pre and post course tests [17]. Modern visual techniques are also used.

At the same time, instruments for assessing spatial ability were also being developed. Already in 1931, the Mental Cutting Test (hereinafter MCT) was developed and in 1971 Shepard and Metzler developed the Mental Rotating Test (hereinafter MRT) [20]. In addition to those two tests, the DAT (Differential Aptitude Test, developed in 1990) and the TPS (Spatial

Imagination Test, developed in 2003) have also been developed. As evident from their names, the first two (MCT and MRT) specialise more in a specific segment of spatial ability (MRT specialises in the spatial ability of rotation and mirroring and MCT in relationships between planes and objects), while DAT and especially TPS are aimed at covering all areas of spatial ability. A common point of all tests is their aspiration for development and improvement. Suzuki and Shiina [21] thus transformed and improved the MRT. What is more, they unified the difficulty of each studied case.

When studying spatial ability, we are usually dealing with a one-off event that attempts to cover a specific population in a specific and locally limited area. Test results are mostly alike and indicate gender differences and advances according to year of study [11]. There are unfortunately not many long-term studies from the field of spatial ability and even these are limited to shorter periods of time, e.g. two years [24]. The reasons for such a situation lie in the substantial number of factors that influence the development of the spatial ability of generations. In order to be able to make a realistic comparison, we also need to have the same testing conditions, which means that different improvements of tests are questionable in this case, regardless of the advantages that they bring. The Faculty of Architecture of the University of Ljubljana (hereinafter the FA), established testing conditions already in 1999 and has been upholding them ever since. This enabled us a realistic comparison between generations.

The tested population of first-year students of the FA is also interesting, as the FA educates the highest number of architects in Slovenia. Interest in enrolment is high, regardless of the negative outlooks for employment. The FA thus holds entrance exams and accepts only the best candidates or candidates with the best spatial ability [8]. We can therefore justifiably assume that students of the FA probably represent the best part of their generation with respect to spa-

tial ability.

A long-term analysis is also important with regard to the fact that the level of knowledge of mathematics and technology of new students is supposedly dropping across Europe. Will this drop also be followed by a drop in the level of spatial ability? The reasons that confirm this thesis can probably be found in the environment in which we live. New information technology has also brought new patterns of behaviour among young people. Traditional games of children and adolescents are becoming sparser, as they are being replaced by digital and virtual games. A good part of traditional games was their development of spatial ability, which was inherent to these games. It is difficult to state what consequences lie in this transition, as on the other hand information technology enables easier perception of space and the relationships within. Results that would show long-term trends in this field are therefore especially important.

In 2010, input data also included information on right-handed and left-handed students. In light of the previously stated, there was supposed to be a difference between the two groups. As claimed by Santrock [18], left-handed persons supposedly have more developed spatial abilities than right-handed persons. The share of left-handed persons among architects is supposedly larger than the share they hold in the entire population.

Based on the previously shown and own experiences, we prepared the following hypotheses:

H1) We wished to verify the generally accepted thesis on better spatial ability among the male population.

H2) As the FA has enrolment restrictions, we accept only students with the best spatial ability, which has already been proven by a study dating back to 2004 [8]. In the past, there have been no major changes in educational content in Slovenia, and we can thus deduce or hypothesise that the level of spatial ability does not change and remains the same every year.

H3) Some studies indicate differences between left-handed and right-handed persons. This was the basis for the hypothesis that the spatial ability of left-handed and of right-handed persons differs.

2. MATERIALS AND METHOD

The MRT test sheets with the key and assessment method were kindly sent to us by Kumiko Shiina from the National Centre for University Entrance Examination in Tokyo (Japan). The conditions of implementing the test remained the same throughout the period and encompassed:

- A4 size test sheets with explanation, test examples and questions. Approximately 10 minutes were always allocated to this part, so that students had enough time to become familiar with how to solve them.
- The main part was the solving of the tasks – two cases of 10 tasks, as in Figure 1. For every task, there were two possible solutions. For every 10 tasks, students had a time limit of 3 minutes.
- The testing took place at the start of the academic year in October as part of tutorials for the Descriptive Geography course.



Figure 1: Sample test task.

Since 1999, we have tested 2,094 students. In 2000, tests were not implemented due to objective reasons. The structure of the participants according to years is evident from Figure 2. When assessing the results, we compared gen-

der differences and from 2010 onwards also left- and right-handedness and the study course (Single Masters Study Programme in Architecture and the First Cycle University Study Programme of Urbanism).

A major change in data processing was implemented in autumn of 2011 due to the aforementioned comparative survey in Budapest and Ljubljana. Instead of the previous recording of the number of achieved points of every student, we began noting their answers in Microsoft Excel. We then used software tools to calculate the results, while such a format also allowed us to assess the changed assessment mode. Such saving of data will also be useful for other analyses (e.g. which examples are “easier”, where can we find the highest number of incorrect answers, etc.) that we may wish to conduct in the future.

In the comparative surveys in Budapest (Ybl Faculty of SzIU) and Ljubljana (UL), students were classified into the following groups:

- Overall – Ybl Faculty of SzIU,
- Men – Ybl Faculty of SzIU,
- Women – Ybl Faculty of SzIU,
- Overall – UL,
- Men – UL,
- Women – UL.

In addition to the aforementioned criteria, the aim of this survey was to establish whether the results correspond to normal distribution.

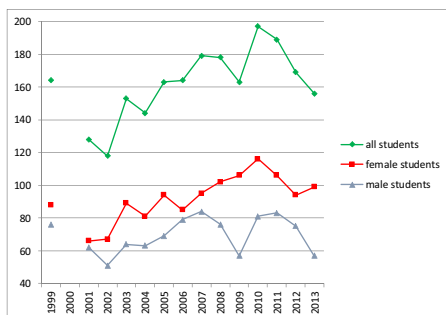


Figure 2: Structure and number of tested students according to study years.

3. RESULTS

3.1 Comparison between the female and male population

The survey of differences between the male and female population encompassed 2,265 persons in the 1999-2013 period. 56.9% were women and 43.1% men. The difference in the average number of achieved points between the two groups is 7.42 and is statistically significant ($P < 0.05$). The results show the anticipated difference according to gender in Table 1.

Table 1: Comparison of differences according to gender.

	women	men
n	1287	978
Average score	17,46542	24,88548
St. deviation	7,187802	8,162033
P	0.0000	

3.2 Comparison of the level of spatial ability in the years 1999-2013

The Figure 3 shows the fluctuation in the average number of points of a generation. The figure also shows that major deviations in the whole population can be seen every three years (2003, 2006, 2009 and 2012). In all these years, the population had a poorer spatial ability. The analysis has statistically confirmed deviations only in the last two cases, i.e. in 2009 ($P < 0.05$) and 2012 ($P < 0.05$). In 2003 ($P = 0.969872$) and 2006 ($P = 0.389689$), the difference was not statistically proven. The figure also shows that sometimes it was more the male population that pushed the results downwards and sometimes the female.

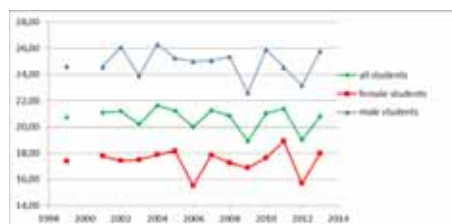


Figure 3: Average number of achieved

points.

The hypothesis on the equal level of spatial ability throughout the period is thus not confirmed. The analysis of the linear trend of increase or decrease of spatial ability shows that spatial ability of students is slowly decreasing, as in Figure 4. This trend is present in all three groups.

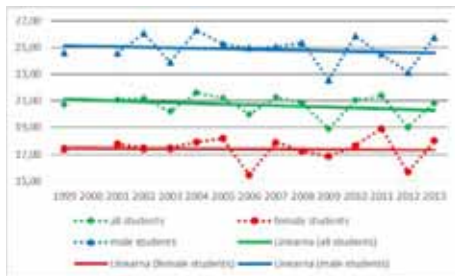


Figure 4: Trends in spatial ability.

3.3 Comparison between left-handed and right-handed persons

The survey encompassed a comparison between all tested students (2010-2013). The difference between right-handed and left-handed persons in the average number of points indicates a 1.08375 point difference ($P = 0.31$) in favour of left-handed persons, which means that there is no statistically proven difference between right-handed and left-handed persons, as in Table 2. It needs to be emphasised that the sample of left-handed persons is small (66 or 10.5% in the entire 2010-2013 period). The hypothesis on differences is not confirmed in this case.

Table 2: Comparison of differences between left-handed and right-handed persons in 2010-2013.

2010-2013	Right-handed	Left-handed
n	629	66
Average	20,50715	21,59091
St. deviation	8,314679	8,417113
P	0,315372	

3.4 Valuation of MRT assessment

A comparative survey was conducted in Budapest and Ljubljana in 2012 and 2013. The survey inter alia indicated that the results do not correspond with the anticipated normal distribution. The reason could also lie in the existing scoring system (2 points are given if both choices are correct, 1 point if only one answer was chosen that was also correct and all other combinations, e.g. one correct and one incorrect, one incorrect, unsolved examples, etc., resulted in 0 points). The analysis of scoring has namely shown certain illogicalities. The number of mixed answers was namely 8 times higher than the anticipated 4:1 ratio. We also believed that a partially correct answer (one correct and one incorrect answer) was better than an incorrect answer (two incorrect answers), even though the existing scoring treats them the same [1]. A new scoring system was thus proposed:

- 2 points for two chosen correct answers
- $\frac{3}{4}$ of a point for one correct and one incorrect answer
- $\frac{5}{4}$ of a point for only one chosen but correct answer
- 0 points for all other choices.

The objective of this comparative analysis was to confirm the hypothesis that with this new scoring system results correspond to normal distribution.

The results of the analysis indicate that with the existing scoring system, only one out of the six groups corresponded to normal distribution, as in Table 3. The implementation of the new scoring system has shown that results correspond to normal distribution to a higher degree, as only one group did not correspond to normal distribution as in Table 4, which confirms the thesis for the need to change the scoring system.

Table 3: Normality test by χ^2 distribution at old scoring system.

Group of students	Computed χ^2 statistic	Answer to H_0 (at 95% $\chi^2 \approx 52.19$)
Overall – Ybl Faculty of SzIU	71.23	NO
Men – Ybl Faculty of SzIU	73.25	NO
Women – Ybl Faculty of SzIU	35.81	NO
Overall – UL	74.61	NO
Men – UL	55.48	NO
Women – UL	68.99	YES

Table 4: Normality test by χ^2 distribution at new scoring system.

Group of students	Computed χ^2 statistic	Answer to H_0 (at 95% $\chi^2 \approx 52.19$)
Overall – Ybl Faculty of SzIU	35.50	YES
Men – Ybl Faculty of SzIU	69.90	NO
Women – Ybl Faculty of SzIU	20.42	YES
Overall – UL	46.50	YES
Men – UL	47.99	YES
Women – UL	34.37	YES

4. DISCUSSION

The survey has shown anticipated gender-related differences and thus confirmed the theory about differences between the male and female population in this field. On the other hand, results do not indicate differences between right-handed and left-handed persons. In light of the already known facts on the differences between left-handed and right-handed persons, which are known from other disciplines, one could expect there to be a difference. The cause may lie in the number of tested persons. The number of tested left-handed persons is nevertheless small, even though the ratio between left-handed and right-handed persons participating in the test does not deviate from the ratio in society. An interesting aspect of the survey is the drop in the level of spatial ability in every third generation. We do not know why this happens. Perhaps there is a similar mechanism as in human biorhythm curves but seen at the student population level over a longer period of time (three years)? Are such fluctuations noticeable also in other fields? There are numerous questions that indicate that this is a new field where more long-term surveys will be needed.

The survey of normal distribution, which was conducted with the help of our colleagues from Budapest, also indicates a need for changing the scoring system. It will be interesting to monitor whether the proposed system will prove to be successful or if additional corrections will be needed and to compare these results with the old system. Modern methods of monitoring and assessment greatly simplify such comparisons.

REFERENCES

- [1] A. Bolcskei, et al. New ideas in scoring the Mental rotating test. Ybl Journal of built environment, Vol. 1., Issue 1, pages 59-69, 2013.
- [2] B. Casey, et al. Use of a storytelling context to improve girls' and boys' geometry skills in kindergarten. Journal of applied devel-

- opmental psychology, 29.
<http://pcs.isiknowledge.com/uml/uml>
- [3] J. P. Gilford. *Personality*. McGraw-Hill, New York, 1996.
- [4] R. Gorska. Modern Research on Spatial Abilities – An Overview and New Results. 11th Scientific and professional Coloquium of CSCGCG: zbornik povzetkov, Varaždinske toplice, 18. – 21. September 2005, Croatian society for geometry and graphics, Zagreb, 2005.
- [5] J. Huttenlocher, N. Newcombe, M. Vasilyeva. Spatial scaling in young children. *Psychological Sciences*, 10(5), pages 393-398, 1999.
- [6] Z. Juscakova, R. Gorska. TPS test development and application into research on spatial abilities. *Journal for geometry and Graphics*, Vol. 11, Nr. 2: 223-237, 2007.
- [7] B. Kožuh. Statistične obdelave v pedagoških raziskavah. *Filozofska fakulteta, Ljubljana*, 2000.
- [8] D. Kušar. Prostorska predstava študentov Fakultete za arhitekturo v Ljubljani. V. *AR Arhitektura, raziskave*: 66-70, 2004.
- [9] M. C. de Lacoste, D. S. Horvath, D. J. Woodward. Possible sex differences in the developing human fetal brain. *Journal of Clinical and Experimental Neuropsychology*, 13, 831-846.
- [10] D. J. Law, J. W. Pellegrino, E. B. Hunt. Comparing the tortoise and the hare: gender differences and experience in dynamic spatial reasoning. *Psychological Science*, 4, 35-40, 1993.
- [11] C. Leopold, R. Gorska, S. A. Sorby. International Experiences in Developing the spatial visualisation abilities of engineering students. *Journal for geometry and Graphics*, Vol. 5, Nr.1: 81-93, 2001.
- [12] M. C. Linn, A. C. Petersen. A meta-analysis of gender differences in spatial ability: Implications for mathematics and science achievement. In J. S. Hyde & M. C. Linn (Eds.), *The psychology of gender: Advances through meta-analysis* (pp. 67-101). Johns Hopkins University Press, Baltimore, 1986.
- [13] M. G. McGee. Human Spatial Abilities: Psychometric studies and environmental, genetic, hormonal and neurological influences, *Psychological Bulletin*, 86, 899-918.
- [14] J. Mohler. A review of spatial ability research. *Engineering Design Graphics Journal*, vol. 72, no. 2, 2008.
- [15] H. Nyborg. Spatial ability in men and women: Review and new theory. *Advances in Behaviour Research and Therapy*, 5(2), 89-140, 1983.
- [16] J. Piaget, B. Inhelder. *Mental imagery in the child*. Basic Books, New York, 1971.
- [17] T. Saito, K. Suzuki, T. Jingu. Relations between spatial ability evaluated by a Mental cutting test and engineering graphics education. V: *Proceedings 8th international conference on engineering computer graphics and Descriptive geometry*, ICGG: 231-235, 1998.
- [18] J. W. Santrock. *Motor, Sensory, and Perceptual Development. A Topical Approach to Life-Span Development* (pgs.172-205) McGraw-Hill Higher Education. Boston, 2008.
- [19] L. Schaik. *Spatial Intelligence*. John Wiley & Sons, Ltd Chichester, 2008.
- [20] R. N. Shepard, J. Metzger. Mental rotation of three-dimensional objects. *Science* 171, 701-703, 1971.
- [21] K. Shiina, K. Suzuki. Design of modified mental rotation test and its error analysis. *Journal for geometry and Graphics*, Vol. 3, Nr. 2: 211-219, 1999.
- [22] K. Shiina, D. R. Short, C. L. Miller, K. Suzuki. Development of software to record solving process of a mental rotating test.

Journal for geometry and Graphics, Vol. 5, Nr. 2: 193-202, 2001.

- [23] C. Soros. Measurement of development of spatial ability of Hungarian engineering students – questions and results. 2nd Croatian Conference on geometry and Graphics: Abstracts, Šibenik, 5. – 9. September 2010, Croatian society for geometry and graphics, Zagreb.
- [24] K. Takeyama, R. Maeguchi, K. Chibana, K. Yoshida. Evaluation of Objective test using a pair of orthographic projections for descriptive geometry education. Journal for geometry and Graphics, Vol. 3, Nr.1: 99-111, 1999.
- [25] E. Tutsumi, H. P. Schrocker, H. Stachel, G. Weiss. Evaluation of students' spatial abilities in Austria and Germany. Journal for geometry and Graphics, Vol. 9, Nr.1: 107-117, 2005.
- [26] B. Yilmaz. On the development and measurement of spatial ability. International Electronic Journal of Elementary Education. Vol. 1, issue 2, March 2009.

mateja.volgemut@fa.uni-lj.si, by fax: +386 61 125 74 14, or through postal address: University of Ljubljana, Faculty of Architecture, Zoisova 12, 1000 Ljubljana, Slovenia.

ABOUT THE AUTHORS

1. Dr. Domen Kušar, Uni. Dipl. Ing. Arch. is an Assistant Professor Descriptive Geometry and General safety at the Faculty of Architecture, University of Ljubljana, Slovenia. His research interests are Spatial ability, Geometry in Architecture, Geometry and Graphics Educations, Safety in architecture. He can be reached by e-mail: domen.kusar@fa.uni-lj.si, by fax: +386 61 125 74 14, or through postal address: University of Ljubljana, Faculty of Architecture, Zoisova 12, 1000 Ljubljana, Slovenia.

2. Mateja Volgemut, Uni. Dipl. Ing. Arch. is a teaching Assistant Descriptive Geometry at the Faculty of Architecture, University of Ljubljana, Slovenia. She is PhD student on Faculty of Civil and Geodetic Engineering, University of Ljubljana on scientific area of spatial planning. She can be reached by e-mail:

THREE-DIMENSIONAL MODELING AND FINITE ELEMENT ANALYSIS OF INVOLUTE HYPERBOLIC ARCH DAM

Tingna Du¹ and Qingqing Mao² and Qilu Du¹

¹Chongqing Jiaotong University, China ² The University of Hong Kong, China

ABSTRACT: Hyperbolic arch dam is being increasingly adopted in hydraulic and hydro-power projects for its outstanding load-carrying capacity. In this paper, a new type of arch dam, namely variable radius involute hyperbolic arch dam is proposed. The arch axis of the dam is developed according to the trajectory of involute. To provide mathematical and mechanical basis for its application in actual projects: new function derivations are proposed with arch ring equations; automatic computing is achieved by using C++ programing, and a 3D entity model of variable radius involute hyperbolic arch dam is constructed, on the basis of which a finite element analysis model is established and results are obtained.

Keywords: Variable radius involute hyperbolic arch dam, arch axis, arch ring curve equation, 3D model, finite element analysis

1. INTRODUCTION

Arch dam is the most complicated and powerful construct in hydraulic and hydro-power projects. The choice of dam's body type depends heavily on the hydrogeology condition of its location. Arch dam requires a V-shaped valley with equal and stable bank slopes, even and sturdy bedrock without major dislocation. Due to the geological restrictions of existing dam body types, such as single center circular arch, multiple center circular arch, parabolic curve arch, logarithmic spiral curvature arch, the study of a new arch dam has become inevitable.

The decision of arch dam's body type usually consists of geometric modeling and the description of horizontal arch ring. The former can be divided into three different types: continuous geometric model, discrete geometric model and mixed geometric model. In actual projects, horizontal arch ring is usually replaced by single curvature with equal thickness. Since design optimization of arch dam is increasingly applied in major and complex projects, the exploration of new line types is required, in both theory and practice.

2. MATHEMATICAL MODELING OF VARIABLE RADIUS INVOLUTE HYPERBOLIC ARCH DAM

2.1 Forming arch axis

A segment on the involute is adopted to represent the axis of the arch, shown in Fig 2.1. In coordinate system XOY , $t_0 = \pi$ is selected as the starting point of research, then progress slowly. Arc CA denotes the right half of arch axis, and its projection on X is AN . $2AN$ is the distance between the surfaces of available bedrock on both banks above crest elevation. According to design code, the radius of arch axis can be calculated using $R_{\text{axis}} = 0.6L_1$, in Fig 2.1, $R_{\text{axis}} = CO$. M_0 is the corresponding point of tangency when $t_0 = \pi$, $CM_0 \perp OM_0$. Segments a_1M_1 , a_2M_2 , a_3M_3 and a_nM_n are tangents at points M_1 , M_2 , M_3 and M_n on base circle. a_1 , a_2 , a_3 and a_4 are intersection points of tangents and involute, the projection of which on AN are b_1 , b_2 , b_3 and b_n . t_1 , t_2 , t_3 and t_n are corresponding values of t at tangency points M_1 , M_2 , M_3 and M_n . δ denotes the absolute value of the difference of segment Nb_n and AN .

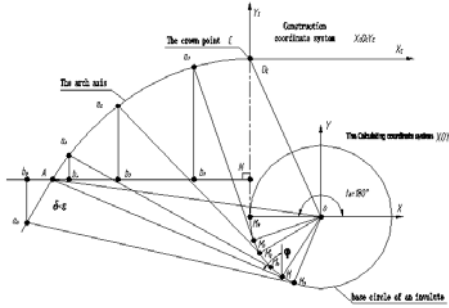


Figure 1: Calculation principle of arch axis and central angle.

In triangle CM_0O :

$$(R_J)^2 + (\pi \cdot R_J)^2 = R_{axis}^2 \quad (2.1)$$

$$R_z \text{ is known, } R_J = \frac{R_{axis}}{\sqrt{1 + \pi^2}} \quad (2.2)$$

In the coordinate system, substituting R_J into the parametric equation of involute:

$$x(t) = R_J(\cos t + t \cdot \sin t) = \frac{R_{axis}}{\sqrt{1 + \pi^2}}(\cos t + t \cdot \sin t) \quad (2.3)$$

$$y(t) = R_J(\sin t - t \cdot \cos t) = \frac{R_{axis}}{\sqrt{1 + \pi^2}}(\sin t - t \cdot \cos t) \quad (2.4)$$

Since:

$$AN + R_J = x(t) \quad (2.5)$$

$$AN = L_1/2 \quad (2.6)$$

Substituting (2.3), (2.4) and (2.6) into (2.5), one has:

$$\frac{R_{axis}}{\sqrt{1 + \pi^2}}(\cos t + t \cdot \sin t) = \frac{L_1}{2} + R_J \quad (2.7)$$

When $t \rightarrow t_n$, $\delta \rightarrow \varepsilon$. As long as the value of t' satisfy equation (2.7), the coordinates of point A in the system can be obtained:

$$x(t') = \frac{R_{axis}}{\sqrt{1 + \pi^2}}(\cos t' + t' \cdot \sin t') \quad (2.8)$$

$$y(t') = \frac{R_{axis}}{\sqrt{1 + \pi^2}}(\sin t' + t' \cdot \cos t') \quad (2.9)$$

Multiple iterations are required during computation to obtain accurate results.

2.2 Solving upstream arch ring curve equation

According to Fig 2.1:

$$x_0 = R_J \quad (2.10)$$

$$y_0 = -\pi \cdot R_J \quad (2.11)$$

According to coordinate transform formula, one has:

$$X(t) = x_0 + x(t) \cdot \cos \theta - y(t) \cdot \sin \theta \quad (2.12)$$

$$Y(t) = y_0 + x(t) \cdot \sin \theta + y(t) \cdot \cos \theta \quad (2.13)$$

Substituting (2.3), (2.10) into (2.12), one has:

$$\begin{aligned} X(t) = & R_J + \frac{R_{axis}}{\sqrt{1 + \pi^2}}(\cos t + t \cdot \sin t) \cdot \cos \theta \\ & - \frac{R_{axis}}{\sqrt{1 + \pi^2}}(\sin t + t \cdot \cos t) \cdot \sin \theta \end{aligned} \quad (2.14)$$

Substituting (2.4), (2.11) into (2.13), one has:

$$\begin{aligned} Y(t) = & -\pi \cdot R_J + \frac{R_{axis}}{\sqrt{1 + \pi^2}}(\cos t + t \cdot \sin t) \cdot \sin \theta \\ & + \frac{R_{axis}}{\sqrt{1 + \pi^2}}(\sin t + t \cdot \cos t) \cdot \cos \theta \end{aligned} \quad (2.15)$$

In the equation: $t \in [t_0, t']$, the value of θ is 0 (when $t = \pi$, two coordinate system are parallel) .

Connecting all the points over the interval, the right half of arch axis is obtained, and the coordinates for the left half can be calculated using the same method. For the convenience of description, arch axis under this circumstance is referred to as $t_0 = \pi$ arch axis.

Upstream curve equation on each elevation can be obtained using the same derivations:

$$X[t][i] = R_j[i] + \frac{R_{axis}[i]}{\sqrt{1+\pi^2}} (\cos t + t \cdot \sin t) \cdot \cos \theta - \frac{R_{axis}[i]}{\sqrt{1+\pi^2}} (\sin t - t \cdot \cos t) \cdot \sin \theta \quad (2.16)$$

$$Y[t][i] = -t_0 \cdot R_j[i] + \frac{R_{axis}[i]}{\sqrt{1+\pi^2}} (\cos t + t \cdot \sin t) \cdot \sin \theta + \frac{R_{axis}[i]}{\sqrt{1+\pi^2}} (\sin t + t \cdot \cos t) \cdot \cos \theta \quad (2.17)$$

In the equation: $t \in [t_0, t'[i]]$

According to (2.16) and (2.17), the corresponding points on the upstream of arch ring can be calculated when value of t_0 changes.

2.3 Solving semi-central angle

As indicated in Fig 2.1, the desired semi-central angle ϕ is the included angle between segment AM and axis y . According to the properties of involute. Normal line AM at point A is tangent

to the base circle of involute, and the length of AM is the radius of curvature at point A . If the slope of AM is known, from inverse tangent function, one has:

$$\tan \phi = -\frac{X(t')}{Y(t')} = -\frac{\cos \theta \cdot \cos t' - \sin \theta \cdot \sin t'}{\cos t' \cdot \sin \theta + \cos \theta \cdot \sin t'}$$

After deformation, inversing function, one has:

$$\phi = \arctan \left(\frac{\sin \theta \cdot \sin t' - \cos \theta \cdot \cos t'}{\cos t' \cdot \sin \theta + \cos \theta \cdot \sin t'} \right)$$

And, semi-central angle on each elevation is:

$$\phi[i] = \arctan \left(\frac{\sin t'[i] \cdot \sin \theta - \cos \theta \cdot \cos t'[i]}{\sin \theta \cdot \cos t'[i] + \cos \theta \cdot \sin t'[i]} \right) \quad (2.18)$$

2.4 Solving downstream arch ring curve equation

Take a point on upstream curve, then make a normal line of the curve from this point and intercept the thickness value at corresponding elevation in the direction of the normal line. The interception point is the desired downstream arch ring curve point. According to advanced mathematics, the parametric equation of space curve Γ is:

$$x = \varphi(t); y = \psi(t);$$

Assuming the two function in the equation are derivable on the interval $[\alpha, \beta]$, and $\omega \in (\alpha, \beta)$, then the equation corresponds a point B $[x_1, y_1]$ on space curve Γ when $t = \omega$, and the tangent equation at point is:

$$\frac{x - x_1}{\varphi'(\omega)} = \frac{y - y_1}{\psi'(\omega)}$$

Assuming a denotes any point on arch ring curve, coordinates $X[t][i], Y[t][i]$. l1 is a tangent involute line passing point a . l2 is the normal line of l1 and tangent to base circle (according to the property of involute). Solving the tangent line equation:

(2.16) and (2.17) are parametric equations of t ,

differentiate, one has:

$$X'[t][i] = \frac{R_{axis}[i]}{\sqrt{1+\pi^2}} (t \cdot \cos t) \cdot \cos \theta - \frac{R_{axis}[i]}{\sqrt{1+\pi^2}} (t \cdot \sin t) \cdot \sin \theta \quad (2.19)$$

$$Y'[t][i] = \frac{R_{axis}[i]}{\sqrt{1+\pi^2}} (t \cdot \cos t) \cdot \sin \theta + \frac{R_{axis}[i]}{\sqrt{1+\pi^2}} (t \cdot \sin t) \cdot \cos \theta \quad (2.20)$$

According to (2.16), (2.17), (2.19) and (2.20), the equation of tangent line passing point a can be obtained:

$$\frac{X_q[t][i] - X[t][i]}{X'[t][i]} = \frac{Y_q[t][i] - Y[t][i]}{Y'[t][i]} \quad (2.21)$$

Slope of tangent line l_1

$$K_q[t][i] = \frac{Y'[t][i]}{X'[t][i]} \quad (2.22)$$

Slope of normal line l_2

$$K_f[t][i] = -\frac{1}{K_q[t][i]} = -\frac{X'[t][i]}{Y'[t][i]} \quad (2.23)$$

According to the coordinates of point a and the slope of normal line l_2 , the equation of downstream arch ring curve at elevation i can be obtained:

$$X_s[t][i] = X[t][i] + \frac{T[i]}{\sqrt{1+K_f[t][i]^2}} \quad (2.24)$$

$$Y_s[t][i] = Y[t][i] + \frac{T[i] \cdot K_f[t][i]}{\sqrt{1+K_f[t][i]^2}} \quad (2.25)$$

Based on the above calculation, the plan of all layers on variable radius involute hyperbolic arch dam is obtained, shown in Fig 2.2. A series of arch rings with different central angle can be generated using this algorithm. Only the plan when $t=\pi$ is given for the sake of concision.

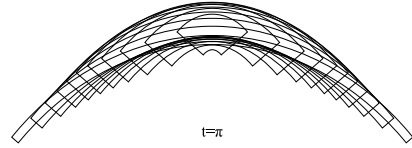


Fig2.2 Arch ring plan when $t_0=\pi$

3. VOLUME CALCULATION OF VARIABLE RADIUS INVOLUTE HYPERBOLIC ARCH DAM AND THE REALIZATION OF 3D ENTITY

Arch ring consists of a series of quadrilaterals lying sequentially, shown in Fig 2.3. In nature, the calculation of the arch ring's surface is the area integral of all the quadrilaterals. The position of downstream arch ring curve is obtained by intercepting the dam's thickness $T[i]$ at corresponding elevation on the normal line of a point on the curve. As a result, when the value of side-length l_{12} (distance between point 1 and 2) of each curve unit is small enough, a quadrilateral approximates a rectangle. The surface value can be obtained by multiplying the thickness of arch dam with the mean value of arch's length sum (both upstream and downstream).

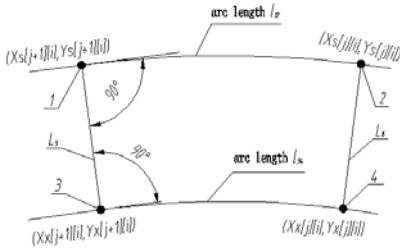


Fig2.3 Solving principle for the small unit area

When the distances between point 1,2 (upstream) and 3,4 (downstream) approach infinitesimal, given L_5 is the normal line of arc l_{12} , and L_6 is the normal line of arc 3,4, then the surface of quadrilateral can be obtained by multiplying the mean length of arc l_{12} and l_{34} with the length of segment L_5 , ie:

$$s_o[j][i] = \frac{l_{12}[j][i] + l_{34}[j][i]}{2} \cdot L_5[j][i];$$

In the equation,

$$l_{12}[j][i] = \sqrt{(X_s[j+1][i] - X_s[j][i])^2 + (Y_s[j+1][i] - Y_s[j][i])^2}$$

$$l_{34}[j][i] = \sqrt{(X_s[j+1][i] - X_s[j][i])^2 + (Y_s[j+1][i] - Y_s[j][i])^2}$$

$$L_5[j][i] = T[j][i]$$

$$V = \lim_{i=0}^N \sum_{i=0}^N \left\{ \frac{\sum_{j=1}^{j=n[i]} s_o[j][i] + \sum_{j=1}^{j=n[i+1]} s_o[j][i+1]}{2} \cdot \Delta h \right\}$$

When difference of elevation $\Delta h \neq 1$, the number of arch ring layers at elevation N is $N/\Delta h + 1$, the volume function of arch dam body V can be obtained.

Setting the difference of elevation small enough,

one has

$$V = \lim_{i=0}^N \sum_{i=0}^N \left\{ \frac{\sum_{j=1}^{j=n[i]} s_o[j][i] + \sum_{j=1}^{j=n[i+1]} s_o[j][i+1]}{2} \cdot \Delta h \right\}$$

When $t_0 = \pi, 2\pi, \dots, 11\pi$, computing the above equations by means of C++ programming, the 3D model of variable radius involute hyperbolic arch dam can be obtained, as shown in Fig 2.4:

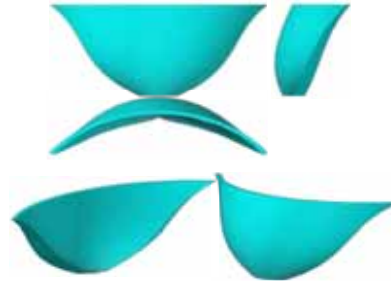


Fig.2.4 Space diagram of involute hyperbolic arch dam

4. FINITE ELEMENT STRESS ANALYSIS VARIABLE RADIUS INVOLUTE HYPERBOLIC ARCH DAM

According to Saint Venant's Principle: If the forces acting on a small portion of the surface of an elastic body are replaced by another statically equivalent system of forces acting on the same portion of the surface, this redistribution of loading produces substantial changes in the stresses locally but has a negligible effect on the stresses at distances. When the foundation of dam is large enough, in comparison with dam body, the stress and strain of the dam body remains the same. Hence, setting the range of finite element model of involute arch dam as: 1.5 dam height towards mountain massif and dam base, ie 150m; exact dam height towards

upstream, ie 100m; 1.5 dam height towards downstream, ie 150m. Taking $t_0 = \pi$ as an example, the computing model of involute arch dam is shown in Fig 2.5.

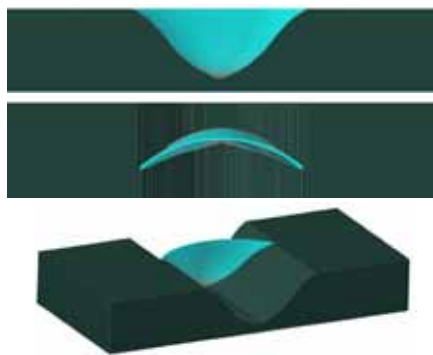


Fig.2.5 Space diagram of arch dam calculation model

Importing the model into ANSYS, and giving each model its deadweight, upstream water pressure, downstream water pressure, silt pressure and temperature load. After calculation and reprocessing, the obtained figures are shown below.

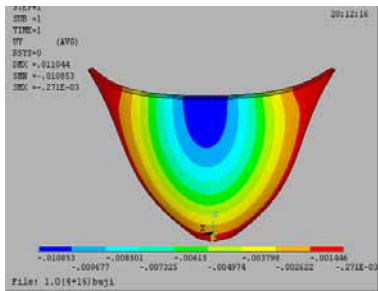


Fig.2.6 Displacement cloud picture in the direction down the river when $t_0 = \pi$

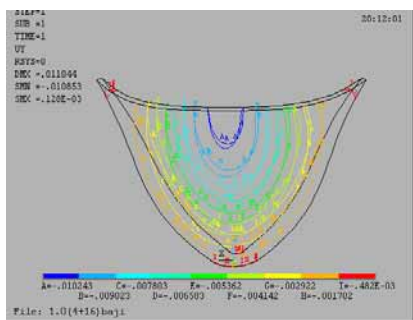


Fig.2.7 Isoline in the direction down the river when $t_0 = \pi$

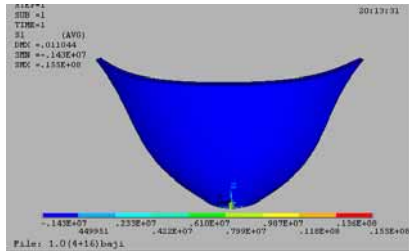


Fig.2.8 The first principal stress cloud picture when $t_0 = \pi$

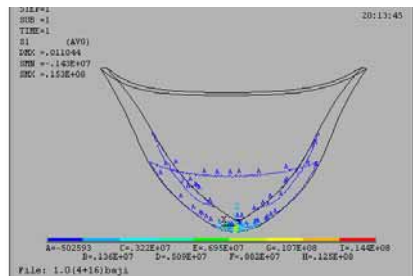


Fig.2.9 Contours of the first principal stress when $t_0 = \pi$

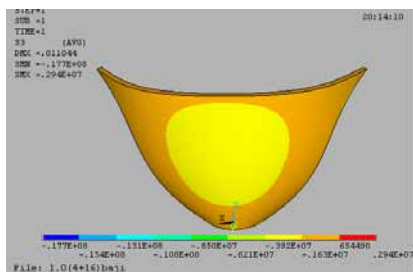


Fig.2.10 The third principal stress cloud picture picture when $t_0=\pi$

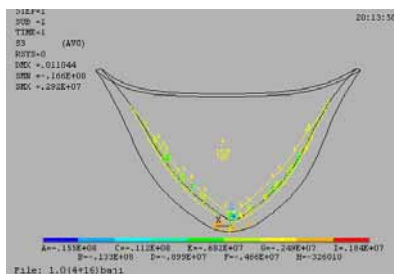


Fig.2.11 Contours of the third principal stress when $t_0=\pi$

5. CONCLUSION

Maximum displacements happen around the crown cantilever of involute arch dam with small differences (Max=10.5mm, Min=10.2mm). When the value t_0 increases, the displacement also increases, but changes slightly, as shown in Fig 2.6 and 2.7. According to Fig 2.8, majority of the first principal stress is compressive stress, as a result of upstream water and silt pressure. The model exhibits stress concentration at dam heel with a maximum tensile stress of 19.8MPa, minimum of 11.2MPa. It is also indicated in Fig 2.9, the maximum crushing stress of upstream and downstream dam surface ranges from 0.45 to 0.61MPa.

According to Fig 2.10 and 2.11, majority of the third principal stress on the dam body is compressive stress (yellow-colored). Yellow zone increases then decreases with the increase of t_0 , and the third principal stress becomes tensile stress with the exhibition of stress concentration. The maximum crushing stress, 2.7MPa, appears at the contacting area between the downstream surface of dam body and bedrock. The above results can serve as a reference for the construction of actual project.

REFERENCES

- [1] Zhong-yuan Jin. *Hydraulic machinery*. Beijing: China hydropower press, 2007.
- [2] Peter Prinz. *The core technology of C language*. Beijing: China Machine Press, 2007.
- [3] Hao-qiang Tan. *The Design of C language procedure*. Beijing: Tsinghua University press, 2002.

ABOUT THE AUTHORS

1. Tingna DU, Female, master of engineering, professor, main research field is computer graphics image processing, the application of CAD technology in water conservancy and hydropower engineering.
2. Qingqing Mao, female, master, graduate student.
3. Qilu Du, male, master, graduate student.

TRADITIONAL DESCRIPTIVE GEOMETRY EDUCATION IN 3D-CAD/CG ERA

Kenjiro SUZUKI

The University of Tokyo (Professor Emeritus), National Institution for Academic Degrees and
University Evaluation, Japan

ABSTRACT: Graphic science is the subject which teaches geometry and graphics, and is taught in early undergraduate curricula at many Japanese universities as a liberal arts subject or as a basic subject for design and drawing. In traditional graphic science courses, descriptive geometry based on manual drawings was taught, but in recent years, the spread throughout society of 3D-CAD/CG has been accompanied by the need for education to respond to these, and graphic science education in the College of Arts and Sciences of the University of Tokyo, where the author taught, 3D-CAD/CG has been implemented to graphic science education since 2007 (Course name: Graphic Science II). With the spread of 3D-CAD/CG, some people say that there is no longer any need for descriptive geometry, but descriptive geometry education is given before 3D-CAD/CG education (Course name: Graphic Science I) as the following reasons; 1) Traditional descriptive geometry is an excellent method in teaching and learning geometry of projection and of three-dimensional objects, and concepts and/or procedures in descriptive geometry can be applied in solving geometric design problems by the use of 3D-CAD/CG. 2) Even in this age of 3D-CAD/CG, hand drawing is still being used (especially for free-hand sketches). 3) Hand drawing is an effective method of developing spatial ability of students. However, with the spread of 3D-CAD/CG, the descriptive geometry techniques in analyzing shapes and forms of three-dimensional objects are now losing their earlier practical importance. So emphasis is not being placed on the education of practical techniques, but is being placed on teaching the theory behind the techniques, i.e., geometry of projection and of three-dimensional objects. In this paper, characteristics features of traditional descriptive geometry, together with specific examples of classes, will be discussed in order to describe the importance of descriptive geometry education, and the need to switch from education focused on techniques to education on the theory behind the techniques.

Keywords: Descriptive geometry, graphics education, graphics literacy.

1. INTRODUCTION

Graphic science is the subject which teaches geometry and graphics, and is taught in early undergraduate curricula at many Japanese universities as a liberal arts subject or as a basic subject for design and drawing. In traditional graphic science courses, descriptive geometry based on manual drawings was taught, but in recent years, the spread throughout society of 3D-CAD/CG has been accompanied by the need for education to respond to these, and graphic science education in the College of Arts and Sciences of the University of Tokyo, where the author taught,

3D-CAD/CG has been implemented to graphic science education since 2007 (Course name: Graphic Science II) [4, 7, 8]. The significance of traditional descriptive geometry, its educational significance in particular, is considered to be great, and descriptive geometry education (Course name: Graphic Science I) is given before 3D-CAD/CG education. But, with the spread of 3D-CAD/CG, the practicality of descriptive geometry as an analysis technique has been declining, and it is necessary to shift educational priority from education focused on practical techniques to the theory behind the techniques, i.e., geometry of projection and of

three-dimensional objects [8, 9].

In this paper, characteristic features of traditional descriptive geometry, together with specific examples of classes, will be discussed in order to describe the importance of descriptive geometry education, and the need to switch from education focused on techniques to education on the theory behind the techniques.

2. WHAT IS DESCRIPTIVE GEOMETRY ?

Before beginning the discussions on specific examples of classes, the characteristic features of traditional descriptive geometry will be clarified, especially from an educational point of view.

2.1 Educational contents of traditional descriptive geometry

Educational contents of traditional descriptive geometry have been divided broadly into two parts. The first is the graphic representation of three-dimensional objects, i.e., projection, and another is the analytical and constitutional method of three-dimensional shapes by the use of their graphic representation.

The term, descriptive geometry, includes the word "geometry", and as shown by the fact that in Europe it is occasionally called "Constructive Geometry", it has been developed into and has been taught as geometry for analysis and construction of three-dimensional objects. Before the emergence of computers, descriptive geometry was the only practical way for analysis and construction of three-dimensional shapes, so education prioritized these techniques.

It is thought that as a result of the long history of technique-centered education, even among people working in descriptive geometry, more have begun to think that descriptive geometry simply deals with techniques. In the background to the techniques of descriptive geometry, there is "Solid Geometry", and it is necessary to be reminded that traditional descriptive geometry education also played an important role as solid geometry education, which is needless to say, important even in 3D-CAD/CG era.

2.2 Characteristics of descriptive geometry

The major characteristic of descriptive geometry

is that it solves problems concerning three-dimensional shapes by reducing them to problems concerning their two-dimensional representations, i.e., "Graphics". A typical method for performing the conversion from three-dimensions to two-dimensions is to use projection. H. Stachel pointed out that a benefit of graphical solution by projection is that preparing a projection from a specified direction simplifies solving the problem, and also mentioned that it aids intuitive grasping of the geometric structure of the problem [5].

We, human beings, live in a three-dimensional world, and capture most of spatial information through our visual system, an "Eye". It is noted here that the eye is, basically, a two-dimensional detector. It uses the two-dimensional projected image on the retina, i.e., graphic representation (Figure 1). It can be said that when we are looking at an actual three-dimensional solid shape, we are seeing it through graphics. It is not too much to say that we recognize three-dimensional world only through graphics. Graphic system is naturally built in human beings. That is why graphics are so natural and, hence, useful for us.

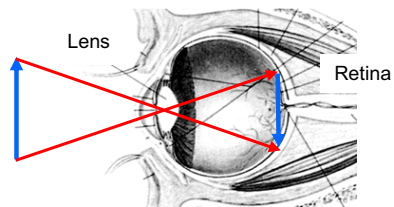


Figure 1: An eye - Two-dimensional detector

Graphics have, however, a fatal defect. This is the fact that because they represent three dimensions in two dimensions, they cannot represent depth direction (projection line direction) information.

It can be said that the essential defect of graphics described above is common to looking at actual three-dimensional objects with our eyes. Let's now imagine village blacksmith who makes horseshoes using ancient methods. A

blacksmith makes horseshoes by hammering the material over and over with a hammer. In this way, he makes three-dimensional shapes. During the work, the blacksmith probably does not vaguely look at the horseshoe. He occasionally rests his hands and looks at the horseshoe from the "front" to make sure that its U-shaped curved state is gradually approaching the shape which conforms with his image, and also, looks at it from the "side" to make sure that is becoming flatter. As mentioned before, the eye is an essentially two-dimensional detector, and it has sharp sensitivity to the two dimensions on the retina surface, but its sensitivity to the depth direction is quite low. To recognize a solid shape correctly, it is not enough to just look at it vaguely, but to look at the shape from a proper direction it can be seen sensitively. It is also necessary to select proper projection direction when we are solving spatial problems through the use of their graphic representations.

Finally, I would like to add one word about the importance of line drawings, which are mainly used in traditional descriptive geometry. Real three-dimensional objects have various attributes - a shape, use, materials, textures, color and so on. It should be noted here that some kind of abstraction is always necessary when we recognize something. Line drawings are the abstracted representation of the shape of actual three-dimensional objects, in which other attributes such as materials, textures and so on are abstracted. Only geometric constitutional relationship of the solid shape, i.e., points, lines and planes are shown in line drawings, which are indispensable for recognizing the shape.

The second characteristic is that the drawings in descriptive geometry are made "manually". H. Stachel has stated that even G. Monge, who established descriptive geometry, did not refer to manual drawing in any way, and that this is not the essence of descriptive geometry [5]. But even now when 3D-CAD/CG has been implemented at not only our college, but many other educational institutions, descriptive geometry is taught with manual drawings. It is widely known

that not only in geometry, but in other types of learning, the concurrent use of sense organs typified by "eyes" and motor organs such as "hands" is a useful way to promote understanding and firmly fix memories. Manual drawing is considered to be important from an educational point of view.

2.3 Why is descriptive geometry taught?

As stated in the introduction, graphic science education at the College of Arts and Sciences of the University of Tokyo includes descriptive geometry education (Course name: Graphics Science I). This is done for the following three reasons.

1) Descriptive geometry is a superior method of teaching and learning geometry of projection and of three-dimensional objects, in particular, for construction and manufacturing, i.e., "Constructive Geometry", since it enables intuitive grasping and rigorous logical reasoning of three-dimensional shapes, and since concepts and/or procedures in descriptive geometry can be applied even in solving geometric design problems by the use of 3D-CAD/CG.

2) Even in the age of 3D-CAD/CG, manual drawing (sketching in particular) is still used and is still practical.

3) Manual drawing is an effective way to enhance spatial ability of students [2, 4].

But with the spread of 3D-CAD/CG, its practicality as an analysis technique based on descriptive geometry (for example, finding the true length of a straight line, finding lines of intersection between two solids) has been declining, and it is now necessary to make a paradigm shift from education with priority on practical techniques to teaching the theory behind the techniques, i.e., geometry of projection and of three-dimensional objects.

3. SPECIFIC EXAMPLES OF EDUCATION

Table 1 shows the typical curriculum of Graphic Science I. It is not very different from that for traditional descriptive geometry with priority on techniques. The differences between lessons

centered on techniques and those centered on the theory behind the techniques are in the details. So this chapter explains differences between the two based on examples of specific teaching of several items.

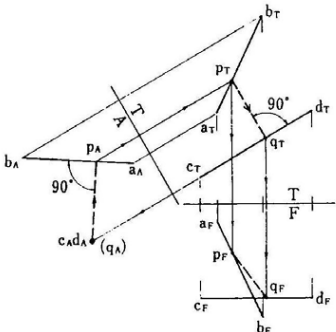
Table 1: Educational contents of GS I
(One 90-minutes time slot per week for 15 weeks)

Week	Contents
1.	Overall guidance
2.	Projection (Principle, Various methods) <i><Analysis and construction of solids by orthographic projection></i>
3.	Principal projection
4.	Auxiliary projection
5-7.	Points, Lines, Planes
8.	Polyhedra (Regular polyhedra / Semi-regular polyhedra)
9-10.	Curved Surfaces (Categorization, / Helical convolute, hyperbolic paraboloid, etc.)
11-13.	Intersections <i><Pictorial projection></i>
14.	Isometric projection / Oblique projection
15.	Perspective projection

3.1 Skew Lines

Figure 2(a) shows a descriptive geometric solution which obtains a straight line connecting two skew lines (AB, CD) at the shortest distance. The minimum distance line of two skew lines is the common perpendicular of the two lines. This is done based on a characteristic of perpendicular projection: "projections of two straight lines intersecting perpendicularly in space intersect perpendicularly, if one of these projections is represented in true length". It is considered to be possible to teach the characteristics of perpendicular projection through these drawings, and at the same time, to inversely, let students understand the geometric characteristics of two skew

lines in space. Figure 2(a) uses a point view and a true length view of straight lines. As already mentioned in Section 2.2, it is only possible to understand the geometric construction of two skew lines, when "they are viewed from a specified direction" such as this.



(a) Descriptive geometry



(b) 3D-CAD

Figure 2: The shortest line which connects two lines in a skew position

It should be noted here that the problem solving procedures in descriptive geometry, i.e., solving problems concerning three-dimensional shapes by reducing them to problems concerning two-dimensional figures, is also useful even in 3D-CAD/CG (see Figure 2(b)), since 3D-CAD/CG uses projective views as an interface with the user, and this is similar to in descriptive geometry [4, 6, 7].

Not only the skew lines shown in this example, but in this course, the explanation is given using

only the auxiliary projection. It does not deal with the "Rabatment" which is a unique technique for manual drawing developed to save space when drawing on paper, and it is presumed to have already become impractical. Also, for similar reasons, it does not deal with "Traces (of a line and a plane)".

3.2 Intersections

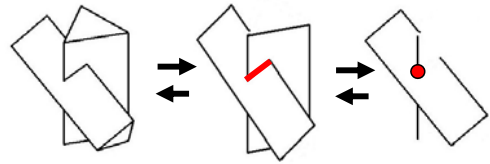
When we create shapes of any kind, we start with a basic solid and remove unnecessary parts - cutting -, or inversely, form the desired shape by combining a number of basic solids - set intersection -. Cutting and set intersection are, geometrically, nothing more than applying the mutual relationships of points, lines, and planes, but they are important as methods for construction and manufacturing.

Figure 3(b) shows a solution method for intersections in descriptive geometry, taking two triangular prisms (A, B) as an example. The solution is as follows;

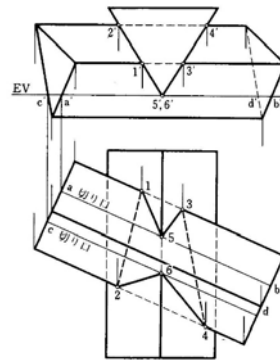
- 1) The points of intersection of the ridge lines of triangular prism (A) and the planes of triangular prism (B) are obtained. The triangular prisms are substituted and the same operation performed, i.e., the intersections of the ridge lines of triangular prism (B) and the planes of triangular prism (A) are obtained. Here, in the example in Figure 3(b), the intersection points (1, 2, 3, 4) are obtained by applying the fact that the side of triangular prism A is represented by an edge view in the top view.

Both triangular prisms are cut on the vertical plane which includes ridge line 5-6 and their sections are used to obtain the intersection points (5, 6). It is also possible to obtain intersection points (5, 6) by creating the edge view of the sides of prism B through an auxiliary projection. But the method of resolving a three-dimensional problem into two dimensions by cutting can be also used in 3D-CAD/CG [6, 7], so cutting method is taught in this descriptive geometry course.

- 2) The line of intersection is obtained by drawing a line connecting the intersection points



(a) Constitutional relationships:
"point \subset line \subset plane \subset solid body"



(b) Intersection between two solids

Figure 3: Intersection

obtained in 1), and an intersection drawing is completed by making a "visible-invisible" judgment. In many descriptive geometry textbooks, the intersection point connecting procedure and/or the "visible-invisible" judgment method are not explained, but are considered to be "imaging a solid." But, there are fixed rules for inducing these from geometric relations [3], and these rules are taught since similar rules are used in 3D-CAD/CG.

It is advocated that before teaching the above solution method, as shown in Figure 3(a), the line of intersection of two solids is formed from intersection lines of "plane and plane" which form each solid, and that the problem of obtaining the intersection lines of plane and plane can be resolved into the problem of finding the point of intersection of "plane and line," and efforts

are made to teach the importance of the geometric constitution: "point \subset line \subset plane \subset solid body." Needless to say, these constitutional relationships are also important in 3D-CAD/CG.

In actual design sites, 3D-CAD/CG is now usually used to practically obtain lines of intersection. But, in order to study the geometric constitutional relationships which are its background, it is important to first manually draw it step by step. It is widely known that not only in geometry, but in other types of learning, the concurrent use of sense organs typified by "eyes" and motor organs such as "hands" is a useful way to promote understanding and firmly fix memories, so it can be said that descriptive geometry is a superior way to teach solid geometry.

Regarding intersections, in addition to intersections of polyhedra cited in this example, intersections of two circular cylinders are taught. A circular cylinder is a curved surface with straight line elements, and considering it in the same way as ridge lines of a polyhedron permits the intersection of curved surfaces to be handled in the same way as polyhedrons. And the intersection of two circular cylinders is also practically important, so this is taught.

In traditional descriptive geometry courses, intersection drawing methods concerning combinations of polyhedrons, cylinders, cones, spheres, and tori in addition to the above examples are also considered. But when 3D-CAD/CG is used, it is possible to obtain solutions far more precisely in a much shorter time, so these solution methods have already lost their practical importance and are no longer taught.

3.3 Regular polyhedra

Construction of regular and semi-regular polyhedra are taught.

Not only in this session, but also in other sessions, application examples are shown as many as possible in order to give the students an understanding of the fact that what they have just studied is not merely a mathematical problem, but is intimately related to real solids (Figure 4).

Showing real application examples may be

also useful to let the students to find "Geometry" from real objects, which have many attributes such as colors, textures and so on. It is no exaggeration to say that nurturing eyes that abstract only shapes from real objects is the ultimate purpose of geometry education.



(a) FCC-lattice
(Tetrahedron +
Octahedron)



(b) Nuclear detection
satellite
(Icosahedron)

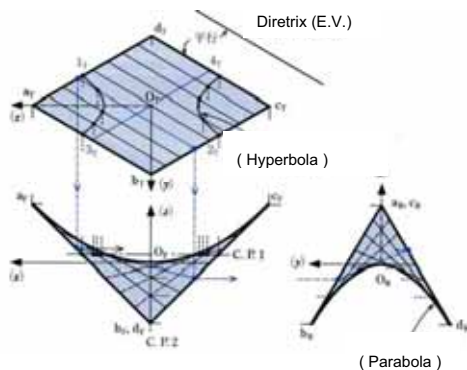
Figure 4: Application of regular-polyhedra

3.4 Curved surfaces

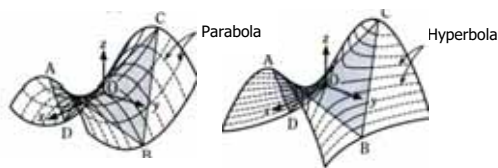
Regarding curved surfaces, the categorization of curved surfaces and some curved surfaces which are important in engineering and architecture, are taught. Categorization of curved surfaces in descriptive geometry is partly based on differential geometry, and is important from the perspective of construction and manufacturing.

The contents of the handling of curved surfaces by descriptive geometry are discussed, taking the hyperbolic paraboloid as an example.

A hyperbolic paraboloid is a curved surface formed when straight line groups parallel to a fixed plane (diretrix) connect two skew lines. As shown in Figure 5(a), a hyperbolic paraboloid has straight line elements and adjoining straight line elements are at the "skew position." In other words, they are not on the same plane. Therefore, this curved surface (theoretically) cannot be developed, and this fact, inversely, means that it is impossible to form a hyperbolic paraboloid from the development diagram without making cuts or wrinkling it. Preparing from a development diagram is often used to process metal sheets, and developability (or non-developability) of a



(a) Straight line elements



(b) Parabolic and hyperbolic elements

Figure 5: Hyperbolic paraboloid



Figure 6: An application of hyperbolic paraboloid (H-P shell) [1]

curved surface is important from the perspective of construction and manufacturing.

A hyperbolic paraboloid has straight line elements. This fact means that it is possible for this curved surface to "be drawn using rulers." But the fact that it has straight line elements means

that it not only can be drawn using rulers, but also it can be made using linear members, lumbers, for example.

In recent years, the advance of 3D-CAD and manufacturing technologies has enabled design and construction of free-form curved surfaces, and these are now used to make automobile and aircraft bodies etc. But, when actually making something, if it is possible to make it with a simpler method, a complex method should not be adopted. Taking the hyperbolic paraboloid described above as an example, it has parabolic or hyperbolic elements (Figure 5(b)), so it is, in principle, possible to construct it by cutting out parabolic shaped or hyperbolic shaped members from a plate and superimposing them. But processing is very difficult, and straight line members, lumbers, are used to construct hyperbolic paraboloid shell (Figure 6).

The discussion in this section has been limited to a hyperbolic paraboloid, but the importance of curved surface education in descriptive geometry, such as helicoid which are widely used in construction of spiral staircase and screws, and helical-convolute surfaces which are applied to gears and to embankments of roads, together with basic concepts and categorizations concerning curved lines or curved surfaces, has not declined at all, even now that 3D-CAD/CG has come into wide use.

4. SUMMARY AND CONCLUSIONS

This paper discusses characteristics features of traditional descriptive geometry, together with specific examples of classes, in order to describe the importance of descriptive geometry education, and the need to switch from education focused on techniques to education on the theory behind the techniques, i.e., geometry of projection and of three-dimensional objects, especially for construction and manufacturing.

ACKNOWLEDGMENTS

The author wishes to give his sincere gratitude to the colleagues of the Department of Computer and Graphic Sciences, College of Arts and Sciences, The University of Tokyo for their daily

discussions on the development of Graphic Science curricula.

REFERENCES

- [1] G. Hood et al. Geometry of engineering drawing. McGraw-Hill, New York, USA, 1969.
- [2] K. Suzuki. Evaluation of student' spatial ability by a mental cutting test - Review on the studies in the past decade. *Proc. 11th ICGG*, 3, 1-5, 2004
- [3] K. Suzuki. A new method in teaching connection of intersection points between two solids (in Japanese). *Proc. 2005 Annual Conference of JSGS*, 103-108, 2005.
- [4] K. Suzuki et al. Development of graphics literacy education (2) - Full implementation at The University of Tokyo in 2007. *Proc. 13th ICGG*, K. Suzuki_I, 1-8, 2008.
- [5] H. Stachel. The study of today's descriptive geometry related education (CAD/CG/DG) in Europe. *Proc. 2007 Annual Conference of JSGS*, 15-20, 2007.
- [6] K. Suzuki and H-P. Schroecker. Application of descriptive geometry procedures in solving spatial problems with feature and parametric modelling 3D-CAD. *Proc. 13th ICGG*, K. Suzuki_II, 1-8, 2008.
- [7] K. Suzuki, Y. Yokoyama and T. Kanai. Introduction to 3D-CAD/CG - Learning geometry and graphics through Inventor and 3ds Max - 2nd Edition (in Japanese). Saiensu, Tokyo, Japan, 2012.
- [8] K. Suzuki. Graphics literacy education at The University of Tokyo. *Series seminar for 150th Anniversary of Politecnico di Milano "The visual language of technique between science and art - Heritage and expectations in research and teaching, No. 3, "Education"*, July 8, 2013.
- [9] K. Suzuki, Significance of traditional geometry in teaching and learning graphic science. *CADDM*, 23-3, 49-55, 2013.

ABOUT THE AUTHOR

Kenjiro Suzuki is a Professor Emeritus of The University of Tokyo and a Professor of National Institution for Academic Degrees and University Evaluation (NIAD). His research interests are in graphic science and university evaluation. He can be reached by e-mail: suzuki-k@niad.ac.jp, by telephone: +81-42-307-1805, or through postal address: NIAD, 1-29-1, Gakuen-Nishimachi, Kodaira-shi, Tokyo, 187-8587 JAPAN.

TRANSLATION SURFACES AND APPLICATIONS

Lidija PLETENAC¹ and Željka MILIN ŠIPUŠ²

¹University of Rijeka, Croatia

²University of Zagreb, Croatia

ABSTRACT: Translation surfaces in Euclidean 3-space E^3 are an interesting topic in the differential geometry of surfaces. The same can be stated for surfaces of constant mean or Gaussian curvature. These classes of surfaces have been explored by many authors and for various purposes. For example, minimal surfaces as well as translation surfaces have an important role in modern civil engineering and architecture. They bring important properties to geometric characteristics of structures (geometry of structure) which are the basis of every construction project.

In this paper, we focus on properties of translation surfaces which are important for their application in architecture and constructions. The aim is a geometric contribution to certain form-finding methods. Translation surfaces are formed by two curves, so-called generating curves, sliding along one another. In this way, a point of the first curve traces out the second curve. They can be classified with respect to properties concerning position and curvature of their generating curves.

Minimal surfaces are surfaces with vanishing mean curvature, $H = 0$. Besides planes, they are non-developable, they minimize the area for fixed boundary conditions and they can be treated as translation surfaces with isotropic generating curves. The only minimal translation surface with both real generating planar curves is the Scherk surface $z = \frac{1}{a} \log(\cos(ax)) - \frac{1}{a} \log(\cos(ay))$, $a \in \mathbb{R} \setminus \{0\}$. Its generating curves are equal resistance catenaries.

Minimal surfaces are theoretically ideal forms for prestressed membrane structures. They are surfaces of negative Gaussian curvature (saddle shaped) and they are important in architectural geometry and in complex shapes of free-form architecture. Ruled surfaces can solve problems such as approximating a desired shape, especially if rulings can be aligned with asymptotic curves of desired surface.

In this paper, we explore the applicability of some higher order algebraic surfaces in engineering. Surfaces containing lines and circles are highlighted. It is possible to find and use a part of the translation algebraic non-ruled quartic surface, whose some properties correspond to demands of desired minimal surface in engineering. Virtual CAD - modeling of shapes (using Rhino) is based on accurate modeling methods, according to the properties of the surface.

Keywords: Translation surface, minimal surface, modeling, application

1. INTRODUCTION

Surfaces that are formed by moving one curve along the other are called *kinematic surfaces*. If those curves are sliding along one another in such a way that a point on the first curve traces out the second curve, a *translation surface* is formed. Translation surfaces have an important role in architecture and civil engineering. In the design of structures such as membranes, domes, cable grids, barrel vaults, tensional integrity struc-

tures, foldable structures etc, important areas are form-finding, modeling and displacement analysis. Process of finding structure shape of equilibrium, with respect to given surface stress state and edge forces or geometrical boundary conditions, is called *form finding*. Form-finding for lightweight structures is a rather demanding procedure. The shape of such a structure is found by various numerical iterative methods, from boundary conditions and prestress, using phys-

ical models, as corrective. Every form finding process starts with one *approximation* of desired shape. The better it is, the easier this process converges. Geometric shape is not the only criterion for optimal form-finding. Uniform distribution of internal forces is another criterion. For membranes (before the external load is applied), this criterion is met when the shape is a minimal surface. In order to find a minimal surface for given boundaries, system of non-linear equations can be solved numerically, for instance, by using software such as *Mathematica*.

In this paper we treat translation and minimal surfaces with the main idea to find a translation surfaces, whose curvature properties and shape correspond to demands of desired surface. Such a surface can be used in the beginning of form-finding process, instead of usually used minimal surface. We'll discuss examples of algebraic and transcendental surfaces that are rarely used in practice, which can be used as an initial approximation to a saddle-shaped lightweight construction. One example will be a quartic patch bounded by circular arcs. Minimal surface, determined by those circular arcs, is very close to that quartic.

Modeling techniques are also based on geometric properties of surfaces. A model is used as initial approximation for numerical optimization of structures. Architectural conceptual design is usually developed using models (physical and virtual). That is why 3D modeling in CAD is implemented in engineering education at universities. At Civil Engineering Faculty in Rijeka, it has been implemented into geometric courses more than 20 years ago and provided as e-learning geometric courses since 2007.

2. TRANSLATION SURFACES

Theory of translation surfaces in Euclidean space E^3 has been developed within the classical differential geometry. Topics of interest include translation surfaces with constant mean curvature or constant Gaussian curvature. It has been generalized to various ambient spaces, as well (e.g. [9]).



Figure 1: Cylindrical surface

Translation surfaces can be classified with respect to properties concerning position and curvature of their generating curves.

Generalized cylinders are translation surfaces with vanishing Gaussian curvature $K = 0$ and constant mean curvature $H \neq 0$. Therefore, they belong to developable surfaces. One of the generating curves must be a line.

Hyperbolic and elliptic paraboloid are doubly curved, hence non-developable translation surfaces, having two parabolas as generating curves. Elliptic paraboloid has $K > 0$ and hyperbolic paraboloid $K < 0$.

Bohemian dome is a doubly curved non-developable quartic surface, generated by two ellipses or circles. The surface consists of areas with $K < 0$ and $K > 0$. Set of points with $K = 0$ forms a curve.

Wavy surface (egg-box surface) can be generated by translation of wavy curve along another curve. Hence it is not developable.

Helicoid is a helical ruled non-developable minimal surface, and a conoid of infinite order. As a translation surface, it can be obtained by translating two helices.

2.1 Bohemian dome

Bohemian dome (*bohemisches Gewölbe*, *dome de Bohème*) is an algebraic quartic translation surface, generated by a circle (or an ellipse) sliding along a circle (or an ellipse). In architecture, the

name *Czech dome* is usually used for a part of a sphere, over a rectangular ground plan ([1, p. 81]) If a circle c_1 of radius b , parallel to the plane xz , is sliding along a circle c_2 of radius a , parallel to the plane yz , the following is the parametrization of Bohemian dome

$$x = a \cos u, y = b \cos v + a \sin u, z = c \sin v$$

for $u, v \in [0, 2\pi)$.

Its implicit equation is given by

$$(x^2 + y^2 + z^2 - a^2 - b^2)^2 = 4(a^2 - x^2)(b^2 - y^2)$$

Bohemian dome is an algebraic surface of fourth degree, i.e. a quartic, see Figure 2. This

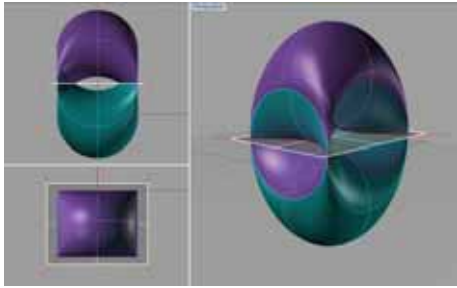


Figure 2: Bohemian dome

surface intersects itself in one of its symmetry planes. The double conic of the surface, having two branches, is a hyperbola.

In the special case when $a = b$, the ground view of the surface is a square and the double curve splits into a pair of perpendicular real double lines. They are asymptotes of the mentioned hyperbola, which is therefore equilateral. These lines are also asymptotic curves on the surface, intersecting in a singular point T with $H = 0$.

Notice that the surface is intersecting itself, which presents a problem for CAD modeling tools (AutoCAD, Rhino ...). However, the whole surface can be modeled by taking a *semicircle* for directrix which is sliding along a generating *circle*, and then reflecting the surface with respect to a symmetry plane in order to get the other half.

2.2 Translation minimal surfaces - a helicoid and the Scherk surface

Considering curvature, a minimal surface is defined as regular surface with vanishing mean curvature $H = 0$. That is the reason why minimal surface can be scaled (from model dimension to the building), keeping the important curvature properties. For given boundary conditions there exists only one minimal surface. Minimal surfaces can be obtained as solutions of Lagrange differential equation. They are non-developable surfaces having minimal area for fixed boundary conditions. Minimal surfaces can be treated as translation surfaces having isotropic generating curves.

The right helicoid is the only ruled minimal surface while catenoid is the only rotational minimal surface. Other examples include the Enneper, Scherk, Catalan and Henneberg surface.

A helicoid. Let α be a helix. We denote by t_α , n_α , b_α its unit tangent, the principal normal and binormal fields. Along α , these fields generate ruled surfaces – a torse, a helicoid and an open helical oblique surface, see Figure 3.

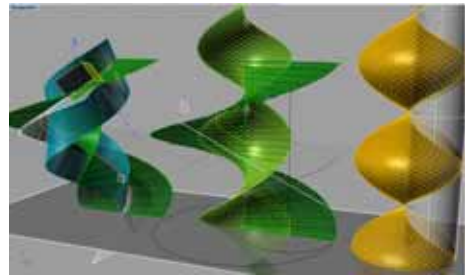


Figure 3: Tangent, the principal normal and the binormal ruled surfaces of a helix

Another interesting helical surface (see Figure 3) is generated as translation surface from two coaxial helices, which are reversely congruent [4, p. 266]. One is the mirror image of the other. Both curves lie on a cylinder which contains the screw axes. This is a surface of revolution and translation at the same time.

Namely, there are three cases in which a translational surface is also a rotational surface: paraboloid of revolution, cylinder of revolution and previously mentioned surface. [5]

The Scherk surface. The famous “chess-board minimal surface” is very known. It is one of Scherk surfaces, the only translation (non-trivial) minimal surface with both real and planar generating curves – they are equal resistance catenaries. The Scherk surface can be parametrized in the form:

$$z = \frac{1}{a} \log(\cos(ax)) - \frac{1}{a} \log(\cos(ay)), \quad a \neq 0$$

Its domain is as a “chess-board”, due to the periodicity of the cosine function.

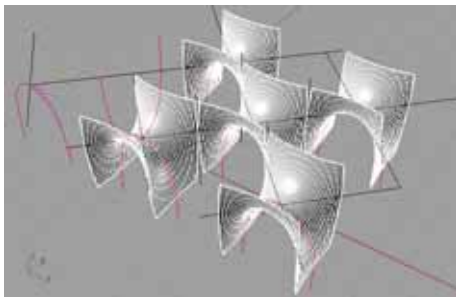


Figure 4: The Scherk surface

Catenary To create CAD 3D model of Scherk surface we need precise model for the generating catenary $z = \frac{1}{a} \log(\cos(ax))$. The equal resistance catenary is formed by weighing a flexible inextensible wire suspended between two points when the linear density (that is to say, in practice, the thickness of the yarn) is proportional to the tension. Its equation is $y = \log \cos ax$, where \log stands for the natural logarithm.

Catenary component in *Grasshopper* computes using \cosh .

For the purpose of this article, catenary was drawn in *Grasshopper* from equation by using cubic interpolation through a chosen number of calculated points.

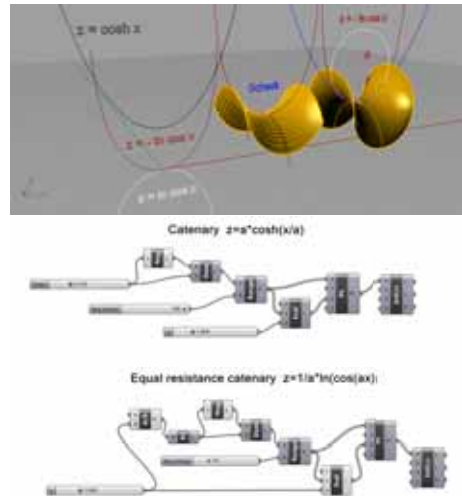


Figure 5: Catenary

3. PROPERTIES OF TRANSLATION SURFACES

Midsurface. Given two directing curves l_1 and l_2 , the set of all lines, connecting each point of the first curve with all points of the second curve is called a congruence [8, 10]. A *translation surface* intersects all rays of a given congruence in midpoints of chords between two directing curves of the congruence.

Proposition. The locus of all chord midpoints, so called midsurface, between two given curves l_1 and l_2 is a translation surface.

Geometric proof. All chords from one point on l_1 intersecting l_2 form a cone with the basis l_2 . Locus of their midpoints is a curve, which is associated to l_2 by central similarity (half the size). As for every point on l_1 there is a cone with the same basis l_2 , there is one family of infinitely many congruent curves $\alpha(u)$, forming surface ψ . The same applies to all points of the other curve l_2 , so there is the other family $\beta(v)$ of curves. Thus, the surface is a translation surface, since it contains two sets of congruent parametric curves.

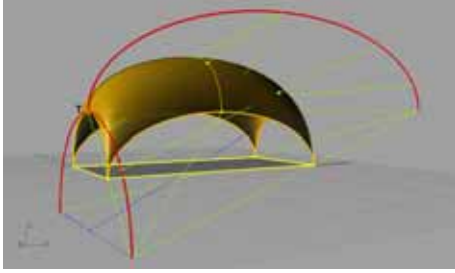


Figure 6: A part of the Bohemian dome as mid-surface

Based on that, one can construct two curves from which the desired translation surfaces ψ will be generated. For the same surface ψ , there are infinitely many pairs of curves (l_1, l_2) . Namely, for any chosen generating curve α there are infinitely many possible positions for l_1 and for each of them there is one curve l_2 .

Example. A part of the Bohemian dome generated as a midsurface is presented in Figure 6. Directing curves of the congruence l_1 and l_2 are constructed by similarity from generating curves (with scale factor 2) and positioned through a chosen point T on surface ϕ .

Example. The right helicoid can also be generated as a midsurface of l_1, l_2 which are both congruent helices ([4]). The midpoint locus of bisecantes of helix h is a curve, similar to h but half sized. Hence, helicoid is translation surface, generated by translating a helix along another congruent helix, as shown in the Figure 3.

Geodesic and asymptotic curves. Geodesic curves are used in designing cable net structures. Asymptotic curves are of interest in free form architecture.

Let $\alpha(u)$ and $\beta(v)$ be generating curves (parametrized by the arc-length) of a translation surface ψ . The translation surface is parametrized by

$$\psi(u, v) = \alpha(u) + \beta(v) \quad (1)$$

Let θ_α (resp. θ_β) be the angle between the surface normal n and the principal normal N_α of the curve α (resp. β). The curvature κ_α , the normal curvature k_α^n and the geodesic curvature k_α^g are related by

$$k_\alpha^n = (k_\alpha^n)^2 + (k_\alpha^g)^2. \quad (2)$$

Furthermore, the following holds ([2])

$$k_\alpha^n = \kappa_\alpha \cos(\theta_\alpha). \quad (3)$$

When $\theta_\alpha = \pi/2$ i.e. $\cos(\theta_\alpha) = 0$ for every point of α , then $k_\alpha^n = 0$ and α is *asymptotic curve* of the surface ψ . Analogously for β . The binormal of the curve is parallel to the normal of the surface.

For example, asymptotic curves of the Bohemian dome (Figure 2) coincide with generating circles in positions $x = \pm b$ and $y = \pm a$. They are locus of parabolic points, with $K = 0$:

On translation surface, Gaussian curvature K vanishes in all points of the generating curve which is asymptotic curve of surface. [2]

If $\theta_\alpha = 0$, i.e. $\cos(\theta_\alpha) = 1$, then $k_\alpha^g = 0$ and α is a *geodesic*. N_α coincide with n . For example, geodesics on the Bohemian dome coincide with generating circles, passing trough the singular point T .

If one generating curve is the asymptotic curve of surface, then it must be a plane curve. [2]

On any regular surface, at every point P with negative Gaussian curvature, there are exactly two asymptotic directions, and they are bisected by principal directions [6]. At a parabolic point there is only one asymptotic direction whereas at a planar point every direction is asymptotic.

Finally notice, that asymptotic directions are perpendicular if and only if the mean curvature H of the surface vanishes at the common point P [6].

3.1 Classes of minimal translation surfaces

Mean curvature of the surface is given by ([2])

$$H = \frac{\kappa_\alpha \cos \theta_\alpha + \kappa_\beta \cos \theta_\beta}{2 \sin^2 \varphi}, \quad (4)$$

According to (3), it can be written:

$$H = \frac{k_{\alpha}^n + k_{\beta}^n}{2 \sin^2 \varphi}. \quad (5)$$

Therefore, a translation surface is minimal ($H = 0$) in the following cases:

1. $\kappa_{\alpha} = \kappa_{\beta} = 0$ This is the trivial case of a minimal translation surface – a plane.
2. $\kappa_{\alpha} = 0, \kappa_{\beta} \neq 0, \cos \theta_{\beta} = 0$. Then α is a straight line, β asymptotic curve, a surface is cylindrical (and minimal), hence a plane.
3. Case for $\kappa_{\alpha} = \kappa_{\beta} \neq 0$ and $\cos \theta_{\alpha} = -\cos \theta_{\beta} \neq 0$. Since the curvatures of α and β are of the different parameter, $\kappa_{\alpha} = \kappa_{\beta}$ holds if and only if $\kappa_{\alpha} = \kappa_{\beta} = \text{const} \neq 0$. Example of such a translation surface is a helicoid whose generating curves are two congruent helices.
4. $k_{\alpha}^n = -k_{\beta}^n$. This is a very general condition, i.e. along α and β a surface is of opposite normal curvature. Example of such a translation surface is the Scherk surface and the Bohemian dome.
5. Case $\kappa_{\alpha} \neq 0, \kappa_{\beta} \neq 0$ and $\cos \theta_{\alpha} = \cos \theta_{\beta} = 0$. Then α and β are planar and they are asymptotic curves of minimal surface. Therefore, α and β are orthogonal.[2]

4. APPLICATIONS OF SURFACES

4.1 Lightweight structures

Lightweight constructions have emerged from the requirement that the ratio of weight and load is as small as possible. Dominant load is snow and wind. According to weight, they can be classified:

1. shells and domes (compressive structures)
2. membranes, freely suspended and prestressed nets (tensile structures)
3. pneumatic (tensile) structures are membranes, inflated with air or helium

Shape is very important for all lightweight structures. They should be smooth, doubly curved surfaces.

Shells and domes. Shells are curved thin structures, which can take load as a membrane. According to Gaussian curvature K , shells are classified into three groups. Geometry of surface is of fundamental importance for behaviour of the structure under load. Compressive stress occurs in shells that have convex curvature. The tension stresses follows the concave curvature. *Translation surfaces* are used for shells in engineering. Wooden structures often have cylindrical shape.



Figure 7: Wooden roof structure

They can also be of other shapes that have Gaussian curvature $K \neq 0$. Free shaped shells can be designed on the basis and with the help of physical models.

Membranes - hyperbolic surfaces. Tensile structures are formatively attractive, very elegant and suitable for covering large areas. They transfer load to their supports through the surface by tensile forces. Because of this, the shape of the surface is of crucial importance and possibly critical property of the structure. Thickness of the membrane is insignificant, self-weight is negligible and materials are flexible (textiles, rubber sheet). Membrane has to be prestressed to avoid slacking caused by wind. Membrane under action of uniform prestress force takes form of a minimal surface. It minimizes the tensile stress in the membrane. Such surfaces exist in nature, hence the name: the natural structures. [7]

There are two types of membranes: tension structures, which need hyperbolic areas of the surface (tent-shaped) and pneumatic structures.



Figure 8: Lightweight construction (Italy) Image taken from Google Street View

Pneumatic structures - elliptic surfaces. As kind of membrane, pneumatic structures are prestressed by internal air pressure in the closed membrane. Double fabric transparent membrane protects from weather and allows the dome to be used in daylight. In nature, strained (closed, inflated) membrane takes the shape of a minimum area.

Cable net structure. Cable net structures and ropes are tensile structures, where the supporting system is made from cables. Minimal cable net is similar to minimal membrane but it is not the same.

In order to simplify calculation of the membrane, one can use an appropriate approximation of the membrane by cable net, which lies on the minimal surface. Boundary conditions of the net, i.e. the coordinates of the boundary nodes are obtained from the boundary curves of surface patch. In case of Bohemian dome, they are circular arcs.

Example: Cable net on Scherk surface gives better results if cables lie on parametric curves (in x and y directions) and make a rectangular grid in the ground plan (x, y). According to [11, p. 56], diagonal grid is not a very good solution. After a number of iterations, cable net is transformed and the result is that forces in elements of net are almost equal.

Advantages and disadvantages of lightweight construction. Cloth is supporting structure but

also the covering. Advantages of lightweight structures are optimally utilized properties (strength) of materials, the possibility of dismantling and recycling, seismic resistance, the way of the force is easily visible etc. Factory production and fast assembly are also advantages. Some of the largest roof structures in the world (like airports and stadiums) prove that the tensile structure are reliable as permanent structures.

Disadvantages of lightweight construction are high cost of labour and materials, low resistance to extreme weather conditions, high complexity and need for high precision in manufacturing. In practice, it is difficult to achieve constant stress in all points of the membrane. Therefore, accuracy of point coordinates for minimal surface is of great importance: Small length variations of the cable can cause major changes of prestressing force in the cable.

Requirements of construction. Stiffness is the resistance to change in shape and dimension. When this resistance comes from the shape of structure, it is called *geometric stiffness*. Tensile structures must be saddled-shaped, so as to be resistant to wind and load. Fibres or cords, forming a network of two systems of curves, resist the gravity and the upward wind action.

Reinforced concrete shells can take the shape of ruled surfaces because the pre-stressing cables can be laid in the direction of surface rulings.

4.2 Form finding processes

The ideal form of *compressive structures* is obtained by *inverse method*: The shape, that took tensile structures under the given load, should be turned upward (reflected with respect to a horizontal plane). By this method structure of a large size can be shaped, minimizing consumption of materials. Optimal form for an arched bridge is catenary. Optimal shape of the dome, under a continuous load is a translation surface: elliptical or rotational paraboloid.

Membrane (tensile structure) is carried by the support structure, which determines the geomet-

ric boundary conditions for membrane structure.

Physical model for tensile structure is soap membrane, which forms the minimal surface, for given boundary conditions.

Nowadays it is appropriate to use geometric CAD models as corrective, within the form finding process.

In the last several years, mathematicians and geometricians developed new algorithms for membrane design. An example of software that uses these algorithms is *K3 Tent*¹. Using this software, tent structure design consists of: tent surface form finding, subdividing the surface into patches and unfolding (approx.) double-curved patches onto the plane, to find cutting pattern. Membrane modeling is possible using software such as *3Ds Max* or *Revit*. *Rhino-Membrane* is a plug-in for *Rhinoceros*, but there is also a plug-in called *Minimal Surface* for *Grasshopper*. Models can be realised by 3D printing. Rapid prototyping is also possible.

5. APPROXIMATIONS OF DESIRED SHAPE

Since Scherk surface is good initial shape for form finding processes, let us compare Bohemian surface with Scherk surface. The question is: Can Bohemian patch be used as first approximation of saddled-shaped structure instead of minimal surface? (see Figure 9).

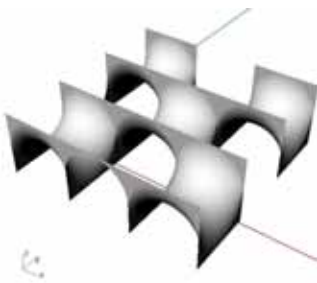


Figure 9: Approximation of the Scherk surface

¹<http://www.k3-tent.com/>

5.1 Approximation by Bohemian dome

Bohemian patch. Let us consider the hyperbolic part of the Bohemian dome as patch. Generating curves are semicircles, forming edges of the patch. Surface lines m and p are horizontal (asymptotic curves at the singular point T).

On some occasions in engineering practice, parabola is not good approximation of catenary. Figure 5 shows that catenary $z = \frac{1}{a} \log(\cos(ax))$ lies above parabola $y = a + \frac{x^2}{2a}$ and below its hyper-osculating circle ($r = \frac{z^2}{a}$) in vertex T (where $r = a$).

Let us create a Bohemian (equilateral) patch with a half of that circle as generating curve. That circle is very close to catenary, as seen on Figure 5. On that figure hyper-osculating circle of parabola is taken as generating arc of Bohemian patch.

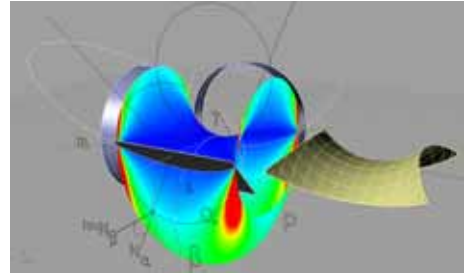


Figure 10: Mean curvature

Using the software *Rhinoceros*, one can check the extreme normal curvatures of the surface, draw osculating circles at each point and check mean and Gaussian curvature. Mean curvature test (Rhino) will show areas where surface comes close to a minimal surface. Choice of the tolerance interval determines how large a part of the area will be marked green, blue or red.

On Figure 10, color indicates mean curvature of the surface: from 0 (blue) to 0.2 (red). Radius of the two generating circles is 10 units.

Mean curvature H vanishes along two lines of surface. It is almost vanishing on the large area of this patch.

Curvature test. Let us compare the mean curvature of the Bohemian patch, a hyperbolic paraboloid and Scherk surface. Mean curvature test (Rhino) will show areas where surface comes close to a minimal surface. Figure 11 shows: a) the Bohemian patch, generated by a circle close to the catenary, b) a hyperbolic paraboloid, generated by parabolas (which can approximate catenary) and c) patch of the Scherk surface, generated by catenaries. A hyperbolic paraboloid is an approximately minimal algebraic surface. It is very popular in architecture and civil engineering, from the last century until now.

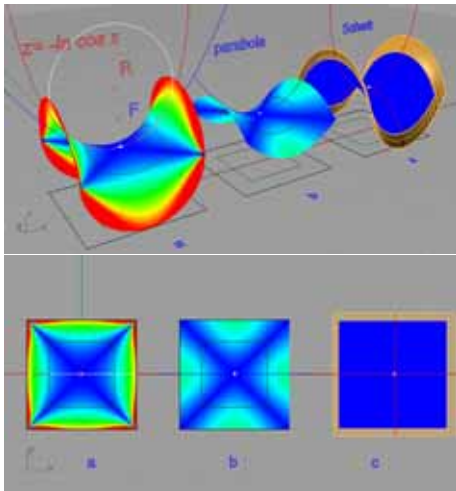


Figure 11: Comparison of surfaces (perspective and top view)

In designing of shallow structures, a parabola is used instead of catenary. Hyperbolic patch of the Bohemian dome can also be an algebraic approximation for minimal surface. In the case of shallow structures it seems to be a better approximation than hyperbolic paraboloid.

Let the suitable part of a regular surface for shallow structures be chosen inside the given boundary, below a (smaller) square in a ground plane (Figure 11). Comparing three mentioned sur-

faces, we can see the difference between approximations by the Bohemian patch and hyperbolic paraboloid:

- Geometric stiffness is better on the Bohemian patch because it is generated by circular arc, which is higher than parabola.
- Covering is better on the Bohemian patch which comes close to the requirement $H = 0$. Mean curvature is very low on the considered region.
- Circular arcs and lines of the Bohemian patch represent facilities in practice (position of supporting elements). For calculations, the coordinates of the boundary nodes can be easily obtained from semicircles.



Figure 12: Soap film

Instead of physical model of minimal surface, one can use CAD model of appropriate accuracy.

The Bohemian hyperbolic patch is not a true unique minimal surface. It satisfies the condition $H = 0$ approximately. The other requirement is that it is stable - i.e. has the minimum surface area between given boundaries. It is interesting to simulate the minimal surface that will appear as a shape of (pre-stressed) membrane between given boundaries (supporting circular arcs). In *Grasshopper*, using iterative algorithm *Rhino-Membrane*, user can try different values in order to have control over design of a surface. As there should be only one solution for a true minimal surface, a physical model of that surface (soap film) would give us the same shape (figure

12). When working with real tensile fabric surface, to get good results, material gravity, tension and stiffness are required.

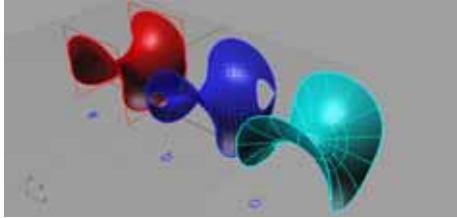


Figure 13: Simulation using *Grasshopper*

For the purpose of this article, that minimal surface was drawn in *Grasshopper* by using *Rhino-Membrane*. Figure 13 shows simulation after cca 100 iterations for boundary arcs as 4 curves (a), joined in two curves (b) and joined in one closed curve (c). Comparison with the Bohemian patch shows that it coincide with minimal surface along surface lines. In other areas minimal surface is a little steeper.

5.2 Approximation by a ruled surface

There are number of papers that deal with asymptotic curves in architectural geometry. For complex shapes in free-form architecture, the technical advantages of ruled surfaces are important. Areas of negative K may be approximated by ruled surface patches. In the initial approximation of desired shape, rulings are aligned with asymptotic curves. Then the shape can be modified to optimally fit a given target shape.[3]

Since a minimal surface has an orthogonal mesh of asymptotic curves, the easiest (and the oldest) way of approximation is use of ruled algebraic quadric surface: hyperbolic paraboloid. Its rulings have to be aligned to asymptotic directions of desired saddled shape. It is an approximation of minimal surface, very often used in engineering.

If the desired saddled shape has to be approximated by *minimal* ruled patch, it is possible to fulfil this demand: Let us consider one of horizontal

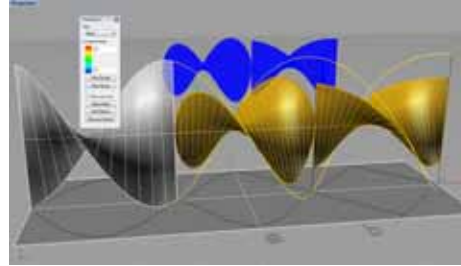


Figure 14: Ruled minimal patch

lines m or p of Bohemian patch (coincident with lines on Scherk surface) as a ruling of a right helicoid, and the other line as the axes of a helicoid. Accurate model is shown on Figure 14. Boundary curves depend on how the minimal surface is trimmed. If surface is trimmed by the sphere (case a), the patch has space boundary curves. If the top view has to be square, surface is trimmed by the prism, boundary curves are plane curves (case b).

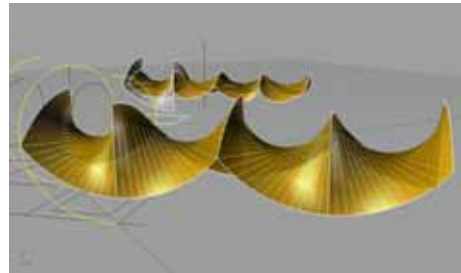


Figure 15: Application of helicoid

Advantages of this saddled patch are: $H = 0$, it is a ruled translation surface, helicoid is stretchable without losing its minimality of surface area. Two horizontal lines m and p on surface (not always perpendicular after stretching) allow us to design supporting structure, forming triangles. Mirroring the patch (above the square top view), one can make modular shape, which could be attractive visually (Figure 16). Since a

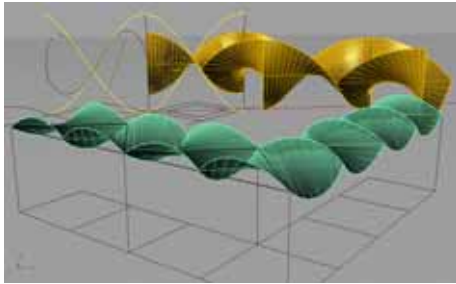


Figure 16: Modular application of helicoid

helicoid is a midsurface of two helices, surface parametrization can be precisely found. In practice surface point coordinates can be precisely found, when measuring on site.

6. CONCLUSIONS

Form finding process for a saddle shaped lightweight structure could start from hyperbolic patch of the equilateral Bohemian dome. Almost in every point of the patch, the mean curvature H is very close to 0. Patch contains two orthogonal lines. Boundary curves are always arcs of the same radius, regardless of how small part of the patch is considered. Minimal surface for that boundaries can always be obtained using *Grasshopper* and *Rhino-Membrane*. Precise coordinates for any point of surface are available from the accurate CAD model or by using properties of mid-surfaces. Furthermore, it is possible to approximate saddled structure by a minimal ruled patch, using right helicoid. A surface also contains two horizontal lines, which may be orthogonal. Namely, surface is stretchable without losing its minimality.

REFERENCES

- [1] H. Brauner and W. Kicking. *Geometrija u graditeljstvu (Baugeometrie)*. Školska knjiga Zagreb, Zagreb, Croatia, 1980.
- [2] M. Cetin, Y. Tuncer, and N. Ekmekci. Translation surfaces in euclidean 3-space. *World*

academy of Science, Engineering and Technology, 5: 757–761, 2011.

- [3] S. Flory and H. Pottmann. Ruled surfaces for rationalization and design in architecture. In *Proceedings of the conference of the association for computer aided design in architecture (ACADIA)*. 2010.
- [4] G. Glaeser. *Geometry and its Applications*. Springer WienNewYork, Austria, 2012.
- [5] G. Glaeser and P. Calvache. On two special classes of surfaces defined by one or more planar or spatial curves. In *Proceedings of the ICGG 15th International conference on geometry and graphics (Montreal, Canada, August 1–5)*. ICGG, Montreal, Canada, 2012.
- [6] A. Gray. *Modern Differential Geometry of Curves and Surfaces*. CRC Press, 1993.
- [7] L. Kalafatić, J. Radić, and Z. Šavor. Analysis of stress-ribbon bridges (proračun prednapetih provješanih sklopova). *Građevinar*, 61(9): 827–836, 2009.
- [8] B. Kučinić, O. Kristoforović, and I. Saler. *Oble forme u graditeljstvu*. Građevinar, Zagreb, Croatia, 1992.
- [9] Ž. Milin Šipuš. Translation surface in a simply isotropic space. *Period. Math. Hung.*, (68): 160–175, 2014.
- [10] V. Niče. *Uvod u sintetičku geometriju*. Školska knjiga, Zagreb, Croatia, 1956.
- [11] R. Vrančić. Proširena metoda gustoća sila. thesis, University of Zagreb, Zagreb, 2013.

ABOUT THE AUTHORS

1. Lidija Pletenac, Faculty of Civil Engineering, University of Rijeka, Croatia.
2. Željka Milin Šipuš, Department of Mathematics, University of Zagreb, Croatia

TRULLIS – ARCHITECTURAL ARCHETYPES

Graciela COLAGRECO¹, Horacio ORTALE², and Gunter WEISS³

¹² University of La Plata, Argentina ³ University of Vienna, Austria

ABSTRACT: The classical trullo seems to fascinate the naïve viewer as well as the educated architect and social anthropologist. The prototypical trulli buildings in Alberobello, Apulia, even became a world cultural heritage. What is so special with trullis? In this paper we use the concept “trullo” for architectural objects with rotational symmetry, ranging from the “Treasure House of Atreus” in Mykene, Hellas, to actual buildings and roof constructions showing rotational symmetry. In this sense “trullo” is an architectural archetype through all epochs and cultures.

There are mainly three viewpoints one might consider, the social and psychological aspect, the architectural and engineering aspect, and finally the geometric point of view aiming at an analysis of how ancient architects defined the meridians of a trullo’s corbelled vault. It seems that there were sort of simple kinematic apparatus in use for that. The paper considers these three aspects, which include also a historical point of view and it continues an article of the same authors, see [3].

Keywords: Trullo, corbelled vault, meridian, kinematics, surfaces of revolution

1. INTRODUCTION

The concept “trullo” is strongly connected with an ensemble of cylindrical buildings with a cone shaped roof and most people think of the village Alberobello in South-Italy, when confronted with the word “trullo”. Figure 1 and 2 show the classical trulli of Alberobello.



Figure 1: Ensemble of Trullis in Alberobello, Apulia, Italy (Source: <http://en.wikipedia.org/wiki/Alberobello>)

Apulia’s trullis have an own and very special history and the local people continue with this building tradition with slight modifications up to now. The main reason for building trullis using rough cobblestone without mortar was that they

quickly could be demolished, when the inhabitants expected a tax collector, such that they could avoided tax demands. Nowadays Apulian trullis are enduring architectural structures, see e.g. Figure 1 and 2.



Figure 2: Trullis – a world cultural heritage
(Source: <http://en.wikipedia.org/wiki/Alberobello>)

But buildings with rotational symmetry exist in every epoch and in every culture, such that it is justified to consider trulli shaped buildings as an architectural archetype comprising also other architectural objects showing rotational symmetry. We are interested especially in the corbelled

vaults of such trulli shaped objects, how they were (and are) designed and finally made. The paper is divided into three main parts: At first we consider the social historic aspect of trulli shaped buildings and what makes a set of architectural objects representing an archetype. A short second part concerns the engineering point of view and it also deals with the question, whether the ancient building techniques used for erecting trullis still might have some relevance or not. The third part deals with the definition of the vault's meridian via primitive kinematic devices. Nothing but curiosity is the problem to find a suitable kinematic device which would allow ancient Greeks to build an existing modern trulli shaped building.

2. ARCHETYPES IN ARCHITECTURE – TOPICS OF SOCIAL-ANTHROPOLOGY AND PSYCHOLOGY

Architecture theory coined the concept of “divine forms in architecture” for a set of four planar figures, namely the *circle*, the *square*, the *equilateral triangle* and the so-called *Golden Rectangle*, the latter representing the proportion of the *Golden Mean* and the regular pentagon, see Figure 3.

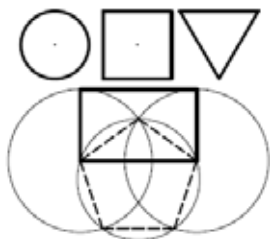


Figure 3: The four fundamental forms of “sacred” architecture

Architecture is three-dimensional, what makes these planar figures so “divine”? The answer could perhaps be found in ancient Egyptian religion, in Greek philosophy and in the origin of Geometry: Egyptian temple architecture had the purpose to “depict” the divine in human reality. How could the transcendent world, its purity and harmony be better visualized than by symmetry

and symmetric objects? No wonder that also the ancient Greeks thought of the real world being a mixture of Platonic solids, which are indeed highly symmetric objects! And no wonder that knowledge in Geometry was a condition sine qua non for being allowed to attend Platon's school of Athens, where people were discussing philosophy. It is therefore more likely that “Geometry” means “inner metric of the transcendent world” than just “land surveying”. (After a usual Nile flood with about 3 mm of mud the borders of each acre surely remained visible. Farmers paid taxes according to their average harvest and not according to the size of their land.) So we state that advanced Architecture and Geometry developed hand in hand as daughters of Religion and Philosophy.

Back to the divine forms in Architecture: In Figure 2 we can see a trullo, the contour of which, in rough abstraction, consists of an equilateral triangle, the contour of the roof, and a Golden Rectangle, the visible part of the cylindrical zone of the building. The entrance shows halves of a square and a circle as gable and tympanum. It seems that special harmonic proportions and a few basic geometric forms and their repetitions in e.g. modules and ornaments evoke similar “groovy feelings” for the naïve viewer, as a harmonic melody does for a listener. Even in early and “primitive” architectural objects, built under the constraints of building material and tools at hand, one can recognize special proportions, ornamentation and some of the “divine archetypical forms”.

One should also mention the psychological and social aspect of „round versus straight“, smooth versus acute, circle versus square or rectangle. Round towers and buildings are eye-catchers and often have the chance to become a town's landmark. The social aspect of a prehistoric group of people with no or only little rank differences sitting around a fire place still remains alive, when we join a “round table meeting”. This idea is still living in the many churches and cathedrals or temples having an apsis: Its center acts as a focus for the representation of the divine. The original earth houses of Sami in North

Scandinavia are “tents” with an “apsis” on two opposite sides. These types of housing seems to date back even to the beginning of mankind. Trulli shaped buildings date back to late Neolithic period and we know examples from that time from excavations in Chirokitia, Cyprus, see Figure 4.



Figure 4: Late Neolithic settlement in Chirokitia, Cyprus. (Source: Ansgar Bovet, http://commons.wikimedia.org/wiki/File:Chirokitia_1.JPG, [2])

3. ENGINEERING ASPECTS OF ANCIENT TRULLIS

Originally *trullis* and the rural *pagghiare* were built of rough cobblestone collected when cleaning an area for agricultural use just like the dry stone walls in the karstic landscapes of Europe, see Figure 5.



Figure 5: Left: Cobblestone fence in Dalmatia (Source: G.Weiss). Right: Cobblestone row material for building a trullo by dry masonry.

Figure 6 shows the flat roofed form of a trullo, which is called *pagghiara* and used as subsidiary building. The image at right shows that the thick

walls of the cone shaped base part consist of an inner and an outer “shell” filled with broken stones in between. The more or less spherical vault combines features of a corbelled vault and a sprung arch, which declares the preference of a conical shape of the base and not a cylindrical basis, see Figure 6.



Figure 6: Pagghiara, a flat roofed trullo, and a collapsed pagghiara near Alberobello. (Source: <http://de.wikipedia.org/wiki/Trullo>, [12])

Figure 7 shows details of the cone shaped roofs of trullis in Alberobello. Here apparently reworked stones were used. Roof constructions by corbelled vaults using rough (or reworked) cobblestone without mortar still ensures relative tightness of the roof, if the support surfaces of the stones have a little outward pointing slope. So even a cupola with radially inclined support surfaces of the bricks would be material saving and had much better static properties people in Apulia kept on building trullis with corbelled vaults up to the 19th century.



Figure 7: Roof of two trullis in Alberobello, detailed view (Source: Mathias Kabel at <http://de.wikipedia.org/wiki/Trullo> [6])

To erect such a stone wall requires skilled and experienced craftsmen and able people in this field are sought after. While an architect mostly is concerned with the (inner) space and purpose of a building and its impression and interaction with its environment, the civil engineer is interested in the material and the methods to perform the building and how stability is ensured. It seems that the perfect rotational symmetry of a trullo and especially its curved meridians made centering sub-structures necessary. Such sub-structures could rotate around the central axis, while a “pointer beam” connected with a simple kinematic apparatus in the meridian plane formed the meridian of the trullo, see Figure 8.

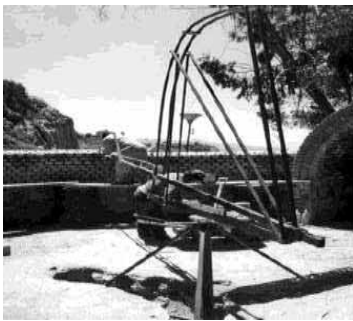


Figure 8: Definition of a trullo's meridian by sort of a gauge setting jig. (Source: <http://en.wikipedia.org/wiki/Trullo>)

The meridian of a trullo's roof is not always a straight line, like it is the case for the trullis in Alberobello. Figures 9 and 10 are medieval examples with a meridian which is close to a conic arc, Figure 11 is a modern trulli with a pear-shaped form and thus a more general meridian. Most of the trullis have rather small windows, if at all, which is due to the building material providing stones with limited size for the window lintel. As references we refer to [1], [10] and [11] and to Wikipedia [12].



Figure 9: Beehive hut, Dingle Peninsula, Co. Kerry, Southern Ireland (Eire). (Source: Dirk Huth at http://de.wikipedia.org/wiki/Kraggewölbebauten_aus_Trockenmauerwerk, [5])



Figure 10: Beehive hut, Bonnieux, Vaucluse (France). (Source: Dominique Repérant at http://de.wikipedia.org/wiki/Kraggewölbebauten_aus_Trockenmauerwerk, [9])



Figure 11: A modern version of a trullo at Flonheim, Germany. (Source: Khjungk at <http://de.wikipedia.org/wiki/Trullo>, [7])

4. GEOMETRIC ANALYSIS OF TRULLI MERIDIANS

The geometric interest lays in the mechanical definition of the meridian of the trullis. The analysis of ancient trullis (e.g. of the “Treasure House of Atreus” in Mykene, (Hellas), Figures 12 - 15) often allows a guess about the used apparatus generating the meridian.

Modern analogues of trulli formed buildings reveal a rich number of rotational surfaces ranging from simple cones of revolution to spindle tori and torus formed surfaces with conic sections but also with splines as meridians. Here the question arises, how the sub-structure Fig.8 had to be modified, if the building would be raised by the ancient methods.

In the following we approach at first to Atreus’ Treasure House.

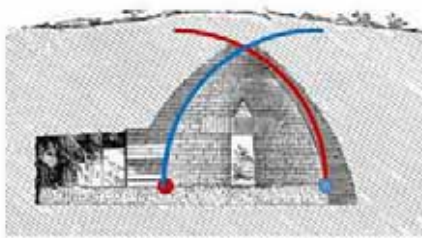


Figure 12: Cross-section of the Treasury House of Atreus, Mykene (Greece). (Source: [13], http://commons.wikimedia.org/wiki/Category:Treasury_of_Atreus?uselang=de#mediaviewer/File:Schatzhaus_des_Atreus,_Querschnitt.jpg)

The drawings of the cross-section reveal slight asymmetries of the meridian. We do not have a 3D-scan at disposal, but it seems that the meridian of this vault is perhaps not exactly a circular arc, c.f. Figure 12 and 13. For a first rough approximation a circular meridian will do revealing an equilateral triangle – a “divine form” - in the cross section. The vault would then be a part of an

ordinary *spindle torus*. It is not clear whether the drawing Figure 12, which stems from 19th century, is based on measurement or rather a schematic visualization. It differs very much from the drawing shown in Figure 13. An approximation by a conic section could be indicated by the cross section Figure 13.

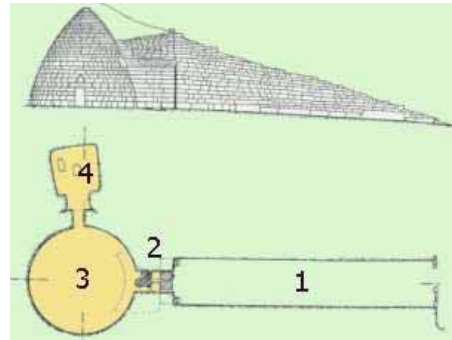


Figure 13: Cross-section and top view of the House of Atreus, Mykene (Greece); 1 Dromos, 2 Stomion, 3 Tholos, 4 burial chamber. (Source: [http://www.gottwein.de/Hell2000/grab001, \[4\]](http://www.gottwein.de/Hell2000/grab001, [4])))

It is also to hypothesize that the vault was built in two sections, the first one up to the architrave of the entrance, a stone of about 120 tons. This might have demanded that the interior had to be filled with broken stones and sand to move that stone to his final place. Then after removing the fill material maybe only partly, a gauge setting jig could have been erected again and the vault then completed. Such a jig surely was a very simple mechanism, c.f. Figure 8. A possibility for a meridian design could be an *ellipse motion*, as shown in Figure 14. It shows a meridian consisting of an arc of an ellipse completed by a circular arc. Conic sections were studied scientifically during the classic period of ancient Hellas, but might have been known as geometric objects already much earlier.

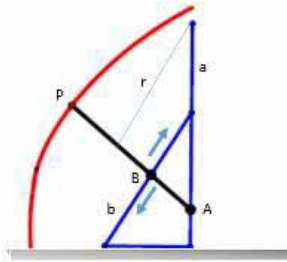


Figure 14: Ellipse motion design as a possible gauge setting jig for Atreus' treasure house.

A rod $[A, P]$ with fixed point B might have been pulled up from the top of an axial beam a by a rope, while A and B slides on a and b . Also modifications of the generated ellipse arc are simple by altering the distances of A , B and P , as well as the slop of b . Such mechanisms can easily rotate around axis a . From Figures 15 and 16 one can see that the used stone bricks are neatly reworked. It is not clear whether the smoothening of the inner surface was done during the building process or after.



Figure 15: Interior view of the Treasury House of Atreus, Mycenae. (Source: <http://commons.wikimedia.org/wiki/User:Ploync?use-lang=de>, [8])



Figure 16: Interior view of the Treasury House of Atreus, Mycenae. (Source: <http://commons.wikimedia.org/wiki/File:Treasure-of-atreus.jpg>, [13])

Another simple mechanism to generate trulli meridians m could be a slider crank type mechanism, see Figure 17: A rod $[A, P]$ with fixed point B on it, which is bound to center C by a rope r , is forced to a planar motion, if A slides along a (horizontal) beam b and the pointer rod slides along the vertical axis a (contact point is Q). Also this mechanism can easily be modified and allows a rich variety of possible meridians.

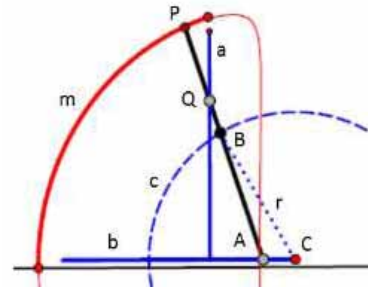


Figure 17: Slider crank motion for trulli meridian design.

The beehive huts Figures 9 and 10 have more or less conic sections as meridians. A possible mechanism for them is shown in Figure 18. It shows the classical generation of an ellipse with points A and B gliding along the axes a and b .

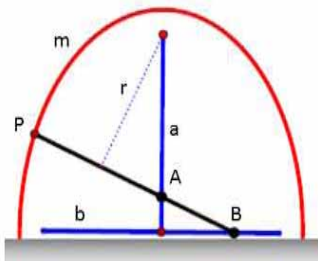


Figure 18: Beehive hut meridian by ellipse motion (see Figure 9 and 10).

A mechanism to design a parabola arc is shown in Figure 19, but as the rod $[P,R]$ sliding through Q orthogonal to a while Q moves along a has fixed length, it is applicable for a trullo from base b only to a limited height. The mechanism uses a rope r of fixed length $\overline{FP} + \overline{PR}$, whereby F acts the focus of the parabola.

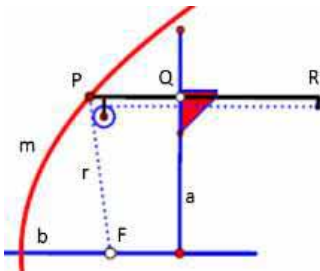


Figure 19: Mechanism to design a parabola arc.

Finally we try to provide a mechanism for the design of the pear shaped meridian of Figure 11. Obviously such a mechanism must be more complicated and it is hardly of practical use. Figure 20 shows a proposal of a coupler mechanism, where the rod points A and B move on circles c and d with centers C and D . The rod point P describes an arc m , which is a suitable approximation of the meridian of the German Flonheim trullo Figure 11. Varying the cranks c and d , the distance of their midpoints C and D and the rod points A , B , P delivers a large set of different shapes of coupler curves.

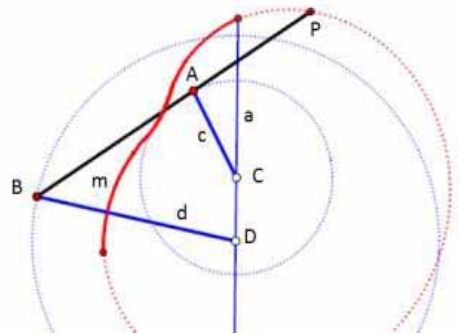


Figure 20: Coupler mechanism to define an arc, which can be related to the trulli meridian Figure 11.

5. CONCLUSIONS

The main concern of this paper is firstly a discussion of the social and historical aspects of trulli buildings. At second we look at the building material and civil engineering aspects of building a trullo. Finally the geometric point of view considers hypothetical mechanisms, which are simple enough such that they could have been used to define the meridian of a trullo.

REFERENCES

- [1] M. Aresta, *Architettura Biologica*, Diseño Editorial, ISBN 978 9873 607264
- [2] A. Bovet, http://commons.wikimedia.org/wiki/File:Chirokitia_1.JPG
- [3] G.M. Colagreco, H. Ortale, G. Weiss, *Las Cónicas en los Trulys*. Proc.VII Intern. Conf. de M&D 2013, Tucuman, Argentina 2013 (to appear.)
- [4] R. Gottwein, <http://www.gottwein.de/Hell2000/grab001>
- [5] D. Huth, http://de.wikipedia.org/wiki/Kraggewölbebauten_aus_Trockenmauerwerk
- [6] M. Kabel, <http://de.wikipedia.org/wiki/Trullo>

- [7] Khjungk, <http://de.wikipedia.org/wiki/Trullo>
- [8] Ploync, <http://commons.wikimedia.org/wiki/User:Ploync?uselang=de>)
- [9] D. Repérant, http://de.wikipedia.org/wiki/Kraggewölbebauten_aus_Trockenmauerwerk
- [10] R. Löbbecke. Kragkuppelbauten. *Verlag der Buchhandlung König*, 2012, ISBN 978-3-86335-100-7.
- [11] B. Spano, La Murgia dei trulli, chapter VII of *La casa rurale nella Puglia*, 1970, p. 184, note 2
- [12] Wikipedia: <http://de.wikipedia.org/wiki/Turullo>
- [13] Wikimedia: http://commons.wikimedia.org/wiki/Category:Treasury_of_Atreus?uselang=de#mediaviewer/File:Schatzhaus_des_Atreus,_Querschnitt.jpg)
- 3. Gunter Weiss, Dr., is a retired professor for Geometry with former employments at the University of Technology Vienna, Austria and Dresden University of Technology, Germany. He can be contacted via e-mail: weissgunter@hotmail.com.

ABOUT THE AUTHORS

1. Graciela Marina Colagreco is a Professor for Physics and Mathematics at the Faculty of Human Sciences and Education at the National University de La Plata. She is director of the Institute for Exact Sciences at this University and also involved in entrance examinations and courses at the Faculty of Medicine and gives Mathematics courses also at the Faculty of Urban Design and Architecture.
2. Horacio Ortale is Surveyor and Constructor Engineer. Ex-Professor for Mathematics at the Faculty of Architecture and Mathematics Analysis at the Faculty of Engineering at the National University of Mar del Plata. Professor for Naval Design at the Faculty of Architecture at the National University of Mar del Plata. Today consultant Professor in Naval Design and co-director at the post grade at the Faculty of Architecture at the National University of Mar del Plata.

TWO POSSIBLE VARIANTS OF GEOMETRICAL MODEL OF CONSTRUCTING KINEMATIC SURFACES ON THE BASE OF INTERIOR REVOLVING ONE AXOID BY ANOTHER ONE

Galina S. RACHKOVSKAYA¹, Yuriy N. KHARABAYEV¹, Natalya S. RACHKOVSKAYA²

¹Rostov State Transport University, Russia

²NMF, USA

ABSTRACT: This paper presents the geometrical model for constructing new kinematic ruled surfaces on the base of interior revolving one axoid by the another one. Kinematic ruled surfaces are constructed by the movement of one of the generating lines of one ruled surface during its movement along another ruled surface. The researching geometrical model corresponds to the case when one axoid is located in the interior of another axoid. As this takes place, there are two possible variants of mutual arrangement of moving and fixed axoids. In the first variant, the moving axoid is located in the interior of the fixed axoid. In the second variant, the fixed axoid is located in the interior of the moving axoid. In the first variant, the outside surface of the moving axoid revolves around the interior surface of the fixed axoid. In the second variant, the interior surface of the moving axoid revolves around the outside surface of the fixed axoid. For such well-known pairs of axoids as “cylinder – cylinder” or “cone – cone” both variants of geometrical model for constructing kinematic surfaces are considered. Visualization of the new constructed kinematic ruled surfaces has been accomplished by the previously developed software application.

Keywords: Analytical Geometry, Computer Graphics, Kinematic Surfaces.

1. INTRODUCTION

Kinematic ruled surfaces can be constructed by the movement of one of the generating lines of one (moving) ruled surface during its movement along another (fixed) ruled surface. The main condition of constructing kinematic ruled surfaces is that moving ruled surface contacts with fixed ruled surface along their common generating line in each of their positions during the movement of one axoid along the other. In some cases such model of the moving as rolling one axoid along another is sufficient to meet this main condition. The examples of these cases are well-known kinematic ruled surfaces constructed on the base of rolling one cylinder along another cylinder or one cone along another cone in the case when one axoid is located outside another axoid [1].

Consider next several options of mutual arrangement of moving and fixed axoids.

Geometrical model **1** corresponds to the case when one axoid is located outside another axoid. In this case, the outside surface of the moving axoid revolves around the outside surface of the fixed axoid (Fig. 1).

Geometrical model **2** corresponds to the case when one axoid is located in the interior of another axoid (Fig. 2). As this takes place, there are two possible variants of mutual arrangement of moving and fixed axoids. In the first variant (A), the moving axoid is located in the interior of the fixed axoid. In the second variant (B), the fixed axoid is located in the interior of the moving axoid. In the first variant, the outside surface of the moving axoid revolves around the interior surface of the fixed axoid. In the second variant, the interior surface of the moving axoid revolves around the outside surface of the fixed axoid.

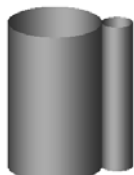


Fig. 1

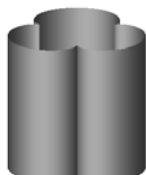


Fig. 1-S



Fig. 2



Fig. 2-S (A)



Fig. 2-S (B)

Contacting cylinders of revolution (Fig. 1, 2) and corresponding kinematic ruled surfaces (Fig. 1-S, 2-S (A), 2-S (B)) are shown.

It should be noted that cases of constructing kinematic ruled surfaces on the base of geometrical model 1 (Fig. 1-S) or on the base of variant A of geometrical model 2 (Fig. 2-S (A)) had been presented in encyclopedic edition [1]. The second variant (B) of geometrical model 2 of constructing kinematic ruled surfaces (Fig. 2-S(B)) has been considered in this paper. Similar geometrical model had been covered before by the example of complex moving one-sheet hyperboloidal surface of revolution along the other [2, 3].

2. GEOMETRIC MODEL OF ROLLING ONE CYLINDER ALONG ANOTHER

Geometrical model of rolling one cylinder of revolution along another (Fig. 1, 2) is presented as a superposition of two interrelated movements: (1) rotational movement of the moving cylinder around its axis (around axis OZ) in the moving coordinate system $OXYZ$ connected with the moving axoid; (2) rotational movement of the axis OZ around the axis oz (lying in axis of the fixed axoid) in the fixed coordinate system $oxyz$ connected with the fixed axoid.

2.1 Geometrical model of outside revolving one cylinder by the another

For the geometrical model 1, parametric representation (in u and v parameters) of kinematic ruled surface, generated by one of the generating lines of the moving cylinder, is based on the following transition equations from the moving coordinate system $OXYZ$ to the fixed coordinate system $oxyz$:

$$x = X \cos u - Y \sin u + (a + b) \cos u;$$

$$y = X \sin u + Y \cos u + (a + b) \sin u;$$

$$z = Z.$$

a – radius of fixed cylinder,

b – radius of moving cylinder.

X, Y, Z are determined by parametric equations of surface, generated by one of the generating lines of the moving cylinder in the moving coordinate system $OXYZ$ [1]:

$$X = b \cos \varphi;$$

$$Y = b \sin \varphi;$$

$$Z = v.$$

The resulting parametric equations of the kinematic ruled surface (Fig. 1-S) in the coordinate system $oxyz$ are:

$$x = b \cos((1 + k)u) - b(1 + k) \cos u;$$

$$y = b \sin((1 + k)u) - b(1 + k) \sin u;$$

$$z = v,$$

where $k = a/b$.

2.2 Geometrical model of interior revolving one cylinder by the another

For the geometrical model 2, transition equations from the moving coordinate system $OXYZ$ to the fixed coordinate system $oxyz$ are:

$$x = X \cos u - Y \sin u + (a - b) \cos u;$$

$$y = X \sin u + Y \cos u + (a - b) \sin u;$$

$$z = Z.$$

For the variant A of geometrical model 2, the resulting parametric equations of the kinematic ruled surface (Fig. 2-S (A)) in the coordinate system $oxyz$ are:

$$x = b \cos((1 - k)u) - b(1 - k) \cos u;$$

$$y = b \sin((1 - k)u) - b(1 - k) \sin u;$$

$$z = v.$$

For the variant B of geometrical model 2, the resulting parametric equations of the kinematic ruled surface (Fig. 2-S (B)) in the coordinate system $oxyz$ are:

$$x = b \cos((1+k)u) - b(1-k) \cos u;$$

$$y = b \sin((1+k)u) - b(1-k) \sin u;$$

$$z = v.$$

3. GEOMETRIC MODEL OF ROLLING ONE CONE ALONG ANOTHER

By analogy with the described above geometrical model of constructing kinematic ruled surfaces on the base of another pair of contacted axoids such as “cone – cone” is considered.



Fig. 3



Fig. 3-S



Fig. 4 (A)



Fig. 4-S (A)



Fig. 5 (B)



Fig. 5-S (B)

Contacting cones of revolution (Fig. 3, 4, 5) and corresponding kinematic ruled surfaces (Fig. 3-S, 4-S, 5-S) are shown.

The axes of the fixed axoids are situated vertically on the Figures 3 - 5.

Geometrical model of rolling one cone of

revolution along another is presented as a superposition of two interrelated movements: (1) rotational movement of the moving cone around its axis (around axis OZ) in the coordinate system $OXYZ$ connected with the moving axoid; (2) rotational movement of the axis OZ (lying in axis of the moving axoid) around the axis oz (lying in axis of the fixed axoid) in the coordinate system $oxyz$ connected with the fixed axoid. Both origins of the coordinate systems $oxyz$ and $OXYZ$ are located in the common vertex of both fixed and moving cones.

For this geometrical model, parametric representation (in u and v parameters) of kinematic ruled surface, generated by one of the generating lines of the moving cone, is based on the following transition equations from the moving system $OXYZ$ to the fixed system $oxyz$:

$$x = X \cos \theta \cos u - Z \sin \theta \cos u - Y \sin u;$$

$$y = X \cos \theta \sin u - Z \sin \theta \sin u + Y \cos u;$$

$$z = X \sin \theta + Z \cos \theta.$$

X, Y, Z are determined by parametric equations of surface, generated by one of the generating lines of the moving cone in the moving coordinate system $OXYZ$ [1]:

$$X = v \sin \beta \cos \varphi;$$

$$Y = v \sin \beta \sin \varphi;$$

$$Z = v \cos \beta.$$

$\theta = \alpha + \beta, \varphi = ku$ are right in the geometrical model 1 (Fig. 3-S).

$\theta = \alpha - \beta, \varphi = -ku$ are right in the variant A of geometrical model 2 (Fig. 4-S (A)).

$\theta = \beta - \alpha, \varphi = ku$ are right in the variant B of geometrical model 2 (Fig. 5-S (B)).

Here α – angle between cone generating line and cone axis for the fixed cone,

β – angle between cone generating line and cone axis for the moving cone,

$$k = \sin \alpha / \sin \beta.$$

Visualization of pairs of contacting axoids and corresponding kinematic ruled surfaces has been accomplished by the previously developed AMG software application [4].

4. CONCLUSIONS

Thus, the new variant of geometrical model for constructing kinematic ruled surfaces on the base of interior revolving one axoid by another axoid in the case when the fixed axoid is located in the interior of the moving axoid is considered. Some new kinematic ruled surfaces are constructed on the base of the researched geometrical model for such well-known pairs of axoids as the pairs “cylinder – cylinder” or “cone – cone”.

REFERENCES

- [1] Krivoshapko, S.N., Ivanov, V.N. Encyclopedia of Analytical Surfaces. M, Nauka. 2010.
- [2] Rachkovskaya, G.S., Kharabayev, Yu.N. Geometric model of kinematic surfaces on the base of one-sheet hyperboloidal surfaces of revolution (one axoid is located in the interior of another axoid). *Proceedings of the 14th International Conference on Geometry and Graphics, Kyoto, Japan*, pages 320-321, 2010.
- [3] Rachkovskaya, G.S., Kharabayev, Yu.N. Geometric modeling and computer graphics of kinematic ruled surfaces on the base of complex moving one axoid along another (one-sheet hyperboloid of revolution as fixed and moving axoids). *Proceedings of the 17-th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision 2009, Plzen, Czech Republic*, pages 31-34, 2009.
- [4] Rachkovskaya, G.S., Kharabayev, Yu.N., and Rachkovskaya, N.S. Computer composition of the transformed classical surfaces as the ways and means of the construction of visual models of realistic objects (The new software application “Art-MathGraph”). *Proceedings of the 15-th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision 2007, Plzen, Czech Republic*, pages 29-32, 2007.

ABOUT THE AUTHORS

1. Galina S. Rachkovskaya, Ph.D., is an associate professor at the Rostov State Transport University, Rostov-on-Don, Russia. Her research interests are Analytic and Differential Geometry of Kinematic Ruled Surfaces, Computer Design of Surfaces. She can be reached by e-mail: g.rachkovskaya@gmail.com or via postal address: 1006 Moorefield Hill Place, Vienna, VA 22180, USA.
2. Yuriy N. Kharabayev is a mathematician-programmer. His research interests are Analytic Geometry, Computer Graphics. He can be reached by e-mail: kharabayev@aanet.ru.
3. Natalya S. Rachkovskaya is a project manager supporting information systems at the Federal Deposit Insurance Corporation, Arlington, VA, USA. She can be reached by email: narachkovska@gmail.com.

USE OF GEOMETRY AT CREATION OR THE ANALYSIS OF DESIGN AND ART OBJECTS

Iryna KUZNETSOVA

National Aviation University (Kiev), Ukraine

ABSTRACT: The article deals with the problem of geometry analysis when creating and using the objects of art and design. Currently, topical in the creation and research of art and design objects is the use of Euclidean, projective, differential, and fractal geometry, the theory of sets, the theory of probability, the theory of dilatation, and geometric transformations. Examples are provided. A simplified classification of geometrical models in industrial design is proposed.

Keywords: design; art; geometry.

1. INTRODUCTION

One of the essential objectives of research in industrial design has been to bring the design problems together within a system of strict and logically consistent scientifically founded knowledge.

2. INFORMATION

The majority of researches are carried out basing on Euclidean geometry, which is naturally accounted for by the character of the researchers' educational background, especially in the case of professional creators of design and art objects. Many existing researches can be extended on the basis of geometry branches that are not traditionally used by engineers and architects.

It is natural to begin with the simplest thing, the point. The perception of dynamic and static compositions of points was described, for example, by N. Yakovlev [5], where Euclidean geometry was made use of.

As an example of priority use of points, we may consider pointillism (from French *pointiller*, i.e. "mark with points")—the method of painting, in which separate distinct brushstrokes (shaped as points or small squares) are applied to the canvas to be optically mixed up in the viewer's eye. The formal and rationalistic character of pointillism, the

fractionality of the piece surface, the locality and brightness of each separate color, did not prevent the use of pointillistic technique in the intensive landscapes by P. Signac with their contrast on colors, or in the canvases by G. Seurat that are full of delicate nuances, or in the portraits and space-and-volume genre compositions by T. van Rysselberghe, etc. A similar perception principle works in other art forms as well, where similar methods are employed based on optical intermixture of point sets, as in the Byzantine mosaic art.

In the course of analysing design and architecture products created by means of points, we here consider points to be indefinable space elements. The space presents itself as a logic form (or structure) providing a medium for implementing other forms. This medium is perceived as a set of objects—points—of any nature.

In geometry, axioms define some external interrelations of elements. Elements form unions of sets. The ways in which such unions are formed, as well as the operations, are structurally limited, in order to avoid paradoxes. Elements are primary to sets. One should note the uniformity of understanding images relative to the understanding of elements and sets. We can consider pointillism in terms of the theory of sets. The concept of

set is one of the broadest concepts in mathematics and in logic. Within the so-called “naive” set theory, the set is defined as an aggregate of any objects which are its constituents and have some characteristic property in common. One can consider sets of points either as an inventory—a list of elements; or as a rule, that defines a particular point’s belonging to a particular image—for example, to a crown of this or that tree, or to the shimmering water surface, or the soft grass on by G. Seurat’s *Sunday Afternoon on the Island of La Grande Jatte* (Figure 1).

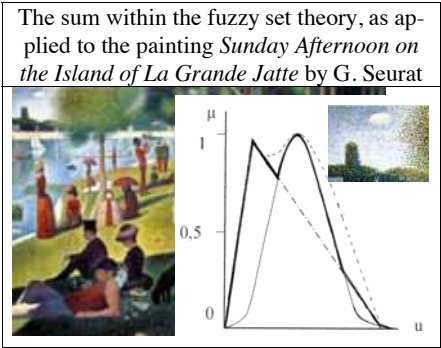


Figure 1

The fuzzy set apparatus appears to be the most acceptable for researching into pointilism.

Following the analysis of point compositions, it is logical to describe the artistic technique of using straight lines in avant-garde painting. This technique is represented by the rayism of M. Larionov, which is considered to be one of the first forms of abstract art and one of the first manifestations of avant-gardism in painting, along with the suprematism of K. Malevich, the futurism of V. Tatlin, and the all-European cubism.

Figure 2 represents an illustration from M. Larionov’s book *The Donkey’s Tail and The Target* and his picture *The Radiant Rooster*. M. Larionov provided the theory of the new artistic trend (1912—1914). Larionov commented

on the “fourth dimension” in art—the radiant and energetic dimension—particularly emphasized by rayism. The idea of multidimensional perception was voiced by many artists, but the multidimensional geometry apparatus appears too complicated to be used for studying art perception.

We may assume though that rayism is based on the apparatus of projective geometry. This geometrical apparatus was also used by other authors of design and art objects. For instance, the whole composition of the *Lord’s Supper* by S. Dali is built on the bunches of projective geometry sheaves.

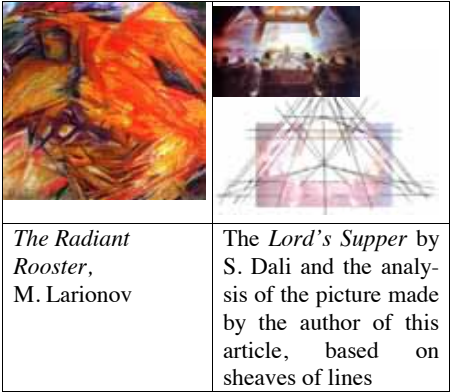


Figure 2

The author of this article has proposed a system of differential equations [1], the first one describing the rotation of the styles’ cyclic recurrence (“linearity” vs “painterliness,” in terms of H. Wöllflin [2]), and the second one reflecting the longitudinal motion of information.

In the general case of a differentiated dynamic system, a phase space is a differentiated manifold which, in turn, is a locally Euclidean space possessing a differential structure. To define a locally Euclidean space we need to correlate it with a Hausdorff topological space. At this stage of research, the author proposes to carry out a comparative analysis of design,

decorative, applied, and graphic arts objects based on the Hausdorff dimension in order to provide a practical solution to the problem of predicting changes in their visual perception.

The second half of the 20th century is characterized by the growing interest in fractal geometry. Fractals are objects with the topological dimension that is smaller than the Housedorff dimension. In art, there exists a fractal ratio known as the Golden Section.



	
1.8814323652 "Linearity," Michelangelo	1.9865897112 "Painterliness," Bernini

Figure 3

Many branches of geometry may be characterised by the type of spaces, and within these spaces, by the type of objects that represent the subject of research.

A number of researches may be pointed out in industrial design, that are connected with explicit or inexplicit use of various sections of geometry.

3. CONCLUSION

At the current stage, the basic techniques of geometrical modeling in creating and studying the objects of design imply the use of

1. Euclidean geometry

2. theory of sets

3. probability theory

1. differential geometry

4. projective geometry

5. fractal geometry

6. dilatation theory

7. geometric transformations.

REFERENCES

- [1] Kuznetsova I. Creating a Set of Differential Equations to Predict Changes of Visual Perception of Design Objects // Proceedings of the 2nd International Conference on Engineering Graphics and Design, June 7-10 2007, Galati, Romania.—pp.106-108.
- [2] H. Wölfflin, Principles of Art History. The Problem of the Development of Style in Later Art. New York, USA: Dover Publication, Inc., 1950. - 237 p.
- [3] Голицин Г.А., Петров В.М. Гармония и алгебра живого. -М.: Знание,1990. -128 p.
- [4] Моль А. Теория информации и эстетическое восприятие. - М.: Мир, 1966. - 678 p.
- [5] Яковлев М.И. Геометричні принципи художнього формоутворення: Дис. д-ра техн. наук: 05.01.03 / КНУБА. - К., 1999.- 352 p.

ABOUT THE AUTHOR

Kuznetsova Iryna, Dokt.Sc., National Aviation University, Prof. of the Department of Interjer Design in Kiev, Ukraine. Her research interest is forecasting, features of perception, formation of the design and art objects form. She can be reached by e-mail: iakuz56@mail.ru

THE USE OF MULTI-MEDIA IN THE TEACHING OF EDUCATIONAL UNIT FOR THE METHODS OF ALTERING SOME PATTERNS OF WOMEN'S CLOTHING TO OVERCOME THE PROBLEM OF FITTING

Thanaa AISARHAN¹ and Hazem ABDELFAH²

¹ King Abdulaziz University, K.S.A ²Helwan University, Egypt

ABSTRACT: Fitting is one of the important criteria for consumers in their buying decision, every garment manufacturer have target segment with certain demographic characteristics, defining consumer profile, so because the fitting issue of garments has impact on the customers buying decision, the main concern of customers especially teenagers and women is to purchase garments that have proper fitting and yet fashionable, although cuts and styles of the clothes we buy is important, the more important factor that influences the buying decision is the proper fitting which is according to the person's body structure, and clothes are not only for body protection and covering, but also have social and emotional aspects attached to them.

For getting the best fit and size dimensions, manufacturer spends big chunks of money, best range of sizing can be a key success factor for manufacturers to implement this many companies are using advanced technologies and strategies to device sizing systems and sizing categories.

There are lots of problems with the fit of commercially made garments, and the manufacturer can avoid that in the step of the sample making by implementing the pattern alterations, which need to experienced pattern maker specially in garment alterations, that inspired the authors with the research idea which aim to allow the student to gain the knowledge and skills allow them to find the best solutions to when, how and where to make pattern altering, adjusting and understanding those pattern modifications.



Figure 1: Some of the Garments Fitting Problems

The authors will design an educational unit to teach the student women's pattern alteration by using

multimedia which can help students acoustically and visually overcome the difficulties and make the concepts and procedures more easily.

The authors will follow the next steps:

1-Determine the general topic of the educational Unit (The Use of Multi-Media in The Teaching of Educational Unit for The Methods of Altering Some Patterns of Women's Clothing to Overcome The Problem of Fitting).

2- Determine the objectives of the educational unit as following:

- The general objectives of the educational unit.
- The procedural objectives of the educational unit.
- After the completion of the application of the unit, the students will be able to achieve the following:
 - A –The knowledge objectives.
 - B– The skill's objectives.

3– Prepare the content of the educational unit.

4 - Evaluate the educational unit by the professors and specialists.

The authors and will apply the educational unit on a representative sample of students from the Department of Clothing and Textile, Faculty of Arts and Designs - King Abdul Aziz University , to determine the effectiveness of the educational unit in the acquisition of knowledge and skills for the students, and will do the following:

- Apply both knowledge and skill tests to measure the knowledge and skills of the students before applying the educational unit and then calculates the results.
- Teach the educational unit using the multimedia acoustically and visually which is explaining how to altering some patterns of women are clothing to overcome some of the fitting problems.
- Apply both knowledge and skills test to measure the knowledge and skills of the students after applying the educational unit and then calculate the results.
- Calculate the significance of differences between the mean scores of the pretest and posttest transactions using appropriate statistical (T test).
- Analyze and discuss the results.

Keywords: Multi-media, Educational Unit, Altering, Patterns, Fitting.

1. INTRODUCTION

In the ever-changing world of education, trends and innovations seem to come and go as often as classes of students. Teachers have little time to adopt new instructional techniques and curriculum before they are outdated and replaced with the “next big thing.” In this fluid environment, one innovation seems to have the potential to become not only a common educational instrument, but one that holds great promise for the future of education as we know it. [5]

The word multimedia means using two or more digital contents in a single application. These contents may be text, video, images, animations, sound, and simulations. This is an area that is constantly evolving and we can say that they are always emerging innovations in the development of multimedia and its application. [12]

As example, Computer programs, hardware, video cameras, microphones and web-servers all play major roles in adding multimedia to a distance course. If facilitators are expected to invest a great deal of time and expense into producing a distance course, they should expect that their efforts will result in an increase in learning and retention by the student when compared to the traditional, text-only version. [13]

This practice is based on the assumption that the use of multiple sensory channels is more effective than the use of each alone. The “cue summation” principle suggests that a combination of channels offers learners a variety of visual and auditory cues, so each learner is able to select the best cues to meet his or her individual sensory needs.

This understanding is confirmed by various research studies and is exercised in the development of any multi-media teaching learning materials such as the study of (Barbara A. Frey and Jann Marie Sutton).

[6]

Using multimedia in our teaching enhance the teaching-learning process—with students’ learning being the focus of all our teaching effort, so the use of computers in education can help students learn more effectively in many ways, and the multimedia courseware allows students high quality instruction independent of time and place. But it can also be integrated with communications software to provide a co-operative and social learning environment. [2]

So multimedia one of the best innovation offers exciting possibilities for meeting the needs of the garment industry, and one of the most important characteristic in this industry is the fitting which is one of the important criteria for consumers in their buying decision and also important for many reasons. Clothing is an extension of oneself, and the way clothing fits affects personal appearance and the first impression one makes on others. Clothing fit also influences how individuals feel about themselves and how they act. Poorly fitting clothing may make a person self-conscious, whereas clothing that fits well allows one to feel self-confident, comfortable, and poised. [11]

Dissatisfaction in the garment industry with fit of women’s wear is widespread in the academic researches such as the studies of (Goldsberry, Shim and Reich) and the results approved that 69 % of a sample of 4,000 women over the age of 55 reported being dissatisfied with fitting in the garment industry. [3]

Also the Studies by Kurt Salmon Associates have consistently shown that more than 50% of the female population cannot find the well fitting garments in the market to fit their bodies, and the women responding to (the Kurt Salmon) Associate’s surveys indicated that fit was the third most frequent reason for not making an garment purchase, which is Consistent with the study of Ashdown, S.P., Delong. [9-8-1]

So fitting varies according to many factors, including

- 1- The style of the garment.
- 2- The fabric being used.
- 3- The end use of the garment.
- 4- Individual preferences.
- 5- Personal size and mood.
- 6- Fashion dictates.
- 7- Religious beliefs of the culture.

And to achieve the right fitting we need to alterate the patterns. [2]

So the preparation of the pattern is one of the most complex processes that the success of the garment depend upon it, which is of the subjects that students have a number of difficulties in the assimilation and in particular the alteration process performed on the Pattern to fit the human body dimension. [7]

The pattern alteration process known as "The process of change, which is performed on the garment's pattern so that it corresponds with the measurements of the body in addition to achieving the right amount of comfort to it"

And the aim of this process is to achieve the well fitting in the garment and this process is known as "Wearing apparel is created when a two/dimensional fabric is shaped through the use of design features such as darts, pleats, tucks, and seems to fit the three-dimensional body form. What is meant by "fit" and how good fit is achieved in personally sewn garments is the focus of this chapter". [4]

For that the researchers thought of using the technological innovations in preparing an educational unit about The use of multimedia in teaching an educational unit for methods of altering some of women's clothing patterns to overcome the problem of fitting which is face a lot of customers and manufactures which also are consistent with the study of (L.J. Anderson, E. L. Brannon, P.V. Ulrich, A.B. Presley , Dave Woronka, M. Grasso and D. Ste-

venson), so the we suggested this educational unit to help the students to enhance their learning process and provide an encouraging educative environment and provide many chances to attract their interest for learning, and to effectively contribute through the presentation vitality and the accuracy, which will help them in concentrating in the information sequence and improve their performance skills . [10]

2. RESEARCH QUESTIONS

2.1 What is the role of the educational institutions in the support and development of the garment industry?

2.2 What are the advantages of the application of the proposed educational unit on the students?

2.3 What is the scientific basis which based on it the educational unit of the knowledge and skills of the students qualify to work in the garment industry?

2.4 What are the components of the educational unit?

2.5 What is the time needed for the education?

2.6 What is the effectiveness of the educational unit proposed in the acquisition of the knowledge contained in the educational unit?

2.7 What is the effectiveness of the educational unit proposed in the acquisition of the basic skills contained in the educational unit?

3. IMPORTANCE OF THE RESEARCH

3.1 Strengthening the educational curriculum.

3.2 The use of multimedia in education gives an element of flexibility in the validity of the teaching of the various categories.

3.3 This research represents a response to the requirements of development and quality of education.

3.4 This research is an educational tool effective for different age groups.

3.5 Highlighting the role of educational institutions in support of the garment industry.

3.6 To contribute to the political part of the Fashion Design Department in the development of curricula in the light of the requirements of the industry.

3.7 support the garment industry through the provision of the proposed unit, which will contribute to the preparation of qualified graduates to work in the garment market.

4. RESEARCH GOALS

4.1 Design of the educational unit for methods of altering some of women's clothing patterns to overcome the problem of fitting.

4.2 Determine the effectiveness of the educational unit in acquisition the knowledge and skills of the students.

5. RESEARCH HYPOTHESES

5.1 There are significant differences between the average scores of students in the knowledge and skills contained in the educational unit pre and post two applications for the post application.

5.2 There are significant differences between the average scores of students in cognitive test before applying the module and then for the benefit of post application.

5.3 There are significant differences between the average scores of students in the skills test before applying the module and then for the benefit of post application.

6. RESEARCH METHODOLOGY

The research follows a quasi-experimental methodology so as to suitability for achieving the objectives of the research and verification of its hypotheses.

7. RESEARCH SAMPLE

The educational unit has been applied to the number of 15 students from the Department of Fashion Design - College of Art and Design - King Abdul Aziz University.

8. RESEARCH TOOLS

8.1 An educational unit for methods of altering some of women's clothing patterns to overcome the problem of fitting.

8.2 A cognitive test (pre / post) to measure the knowledge contained in the educational unit.

8.3 A skill test (pre / post) to measure the skills contained in the educational unit.

8.4 Scale of assessment to evaluate the patterns done by the students.

9. RESEARCH LIMITS

9.1 Measure the effectiveness of the proposed educational unit on a sample of students in Level 8 - department of fashion design - College of Arts and Designs - Jeddah - Saudi Arabia - during the first semester of the academic year (2013/2014).

9.2 Evaluate the educational unit through the application of cognitive achievement test and skills performance test (pre/post).

9.3 The achievement test limited to measure levels of remembering - understanding - the application.

9.4 The skill performance test limited to measure the levels of skill performance to altering some of women's clothing patterns to overcome the problem of fitting.

9.5 The cognitive side has been taught in one of the college halls, also the skill side has been taught in one of the college halls laboratories.

10. RESEARCH PROCEDURES

10.1 Preparation of the educational unit

The researchers prepared an educational unit for teaching methods to modify some styles of women's clothing to overcome the problem of in accordance the right scientific methodology, and the following steps for setting up the unit:

10.1.1 Select a unit topic

Researchers chose the subject of the unit in response to the state's policy in advancing human development in various fields, especially in the area of the garment industry, in order to give students the knowledge and skills to be an fitted constructivism unit from which to make their way to work in the garment industry.

10.1.2 Determine the objective of the educational unit

After the student study the educational unit, will be able to:

10.1.2.1 Be familiar with the concepts and terminology used in the educational unit.

10.1.2.2 Recognize the process of fitting. Mention the methods of fitting and the factors influencing them.

10.1.2.3 Supplied with the necessary basic information in the major of some of the women's types alteration.

10.1.2.4 Show some specific common problems of some of the women's types.

10.1.2.5 Solve specific problems in the fitting.

10.1.3 The period of application

The period of application needed 5 weeks, one day per week, five hours per day.

10.1.4 Place of the educational unit application

The teaching done in the department of fashion design labs - College of Arts and Designs – king abdulaziz university.

10.1.5 The hardware and tools used in the teaching of the unit

10.1.5.1 Tables for drawing and cutting the patterns.

10.1.5.2 Pencils - ruler 30 cm, 60 cm - glue - average size scissor - measure tape.

10.1.5.3 Transparent and colored paper.

10.1.6 Educational tools used to teach the unit

10.1.6.1 Display for the data device (Data Show) to explain some of the lessons.

10.1.6.2 Blackboard.

10.1.6.3 boards to clarify the parts of the altered patterns.

10.1.7 Evaluating the results of the students.

Results of the students are evaluated by:

10.1.7.1 A cognitive test (pre / post) to measure the knowledge contained in the educational unit.

10.1.7.2 A skill test (pre / post) to measure the skills contained in the educational unit.

10.1.7.3 Scale of assessment to evaluate the patterns done by the students.

10.1.8 Validity and Reliability of the Research Tools

10.1.8.1 Validity of the Cognitive Achievement Test

The subject of the test validity as measured by the test and the extent to which the measured succeed .

The Logical Validity

The achievement test has been presented on a jury of professors who specialize in order to ascertain the extent ease and clarity of the test statements, and the objectives linking with the questions test, and the arbitrators were unanimous on the validity of the achievement test for application with to make some proposals, and the followings has been modified based on their Proposals:

- Reduce the number of the questions .
- Take care of the ease and clarity of the wording.

10.1.8.2 Reliability of the Cognitive Achievement Test

Reliability is defined that the test have to be coordinated with what it gives of results, the reliability coefficient achievement test have been calculated in the following methods :

- **Reliability using spilt-half reliability**

It was ascertained the reliability of cognitive achievement test method using spilt-half reliability, and it is an value of the reliability coefficients 0.826 - 0.937, which is significant value at the level of 0.01 per approached from the right one , which shows the reliability of cognitive achievement test .

- **The reliability of alpha coefficient**

Found that the value of alpha coefficient is 0.888, which is high value and this is evidence of the reliability of the achievement test at the level of 0.01 per approached from the right one.

Table 1: The reliability of the cognitive test

the reliability of the cognitive test	Split-half		Cronbach's Alpha	
	Correlations	Sig	Correlations	Sig
	– 0.826 0.937	0.01	0.888	0.01

10.1.9 Validity and reliability of the skills applied test

10.1.9.1 The Validity

The logical Validity

The test was presented to a group of professors and all specialists acknowledged that it all for application.

10.1.9.2 The Reliability

The reliability of scorer

Can be obtained on the reliability coefficient scorer by calculating the correlation

coefficient between the grades given by two scorer or more for the same individuals or for the same tests , in other words , all unexamined gets two grades or more from the correction of one test.

The corrections done by three professors of arbitrators and using an evaluation form in the evaluation process and every corrector did the evaluation process individually.

The correlation coefficient has been calculated between the three grades which done by the correctors (x, y, z) for the post apply test by using the rank correlation coefficient and the following table show this:

Table 2: Correlation coefficient between the correctors for the skills test

The correctors	The bodice	The sleeve	The skirt	The trouser	Total
X,Y	0.816	0.723	0.880	0.922	0.829
X,Z	0.877	0.855	0.913	0.783	0.741
Y,Z	0.759	0.937	0.802	0.861	0.899

Seen from the table above increase in the values of correlation coefficients between the correctors, and all values are a significant at the level of 0.01 per approached from the right one, which shows the reliability of the practical test, which measures the skills performance, and also demonstrates the reliability of the evaluation form and it is tool of the skill test.

11. THE RESULTS

11.1 The First Hypothesis

The first hypothesis state as follows :

“There are significant differences between the average scores of students in the cognitive and skills contained in the educational unit in both pre and post applications for the post application ”.

To investigate this hypothesis has been

applied test "T" and the following table explains this :

Table 3: Significant of the differences between the average scores of the students in the pre and post application for the educational unit

The effec- tiveness	Mean (m)	Std. De- viation	N	df	t	Sig
pre	18.242	6.339	14	15	36.55	0.01
post	118.527	11.182				

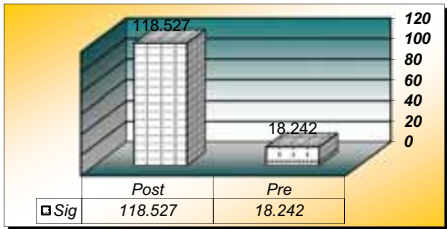


Figure 2: Explain the Differences between the Average of the Student's Scores in the Pre and Post Application for the Educational Unit

Both of table (3) and figure (2) explain that the value of "T" equal "36.554", and it is a value statistically significant at the level of 0.01, where the average of the student's scores of students in the post application is "118.527", while the average of student's scores in the pre application is "18.242", which indicates the existence of real differences between the two applications for the post application, that is, the use of multimedia in teaching an educational unit for methods of altering some of women's clothing patterns to overcome the problem of fitting in this study successful in achieving the goal of it

and already teach the fundamentals contained in it for the cognitive and skills. And to know the magnitude of the impact we applied ETA

Equation :(1)

$t=36.554$

freedom scores =14

$$n^2 = \frac{t^2}{t^2 + df} = 0.989 \quad (1)$$

And by calculate the magnitude of the impact, it found that

$$n^2 = 0.989$$

Equation :(2)

$$d = \frac{2 \sqrt{n^2}}{\sqrt{1-n^2}} = 19.03 \quad (2)$$

The magnitude of the impact is determined by whether big, medium or small, as follows :

Small impact magnitude= 0.2

Medium impact magnitude= 0.5

Large impact magnitude= 0.8

This means that a large effect size, thereby achieving a first hypothesis.

11.2 The Second Hypothesis

The second hypothesis state as follows:

"There are significant differences between the average scores of students in the cognitive test pre and post of the educational unit application for the post application".

To investigate this hypothesis has been

applied test "T" and the following table explains this:

Table 4: Significant of The Differences Between The Average Scores of The Students in The Pre and Post Application for The Cognitive Test

The cognitive test	Mean (m)	Std. Deviation	N	df	t	Sig
pre	1.631	0.721	15	14	10.369	0.01 for the post
post	10.544	1.252				

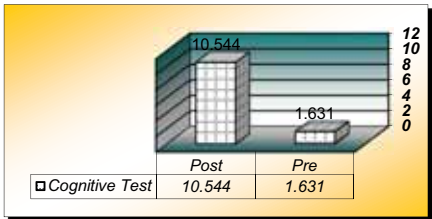


Figure 3: Explain The Differences between the Average of The Student's Scores in The Pre and Post Application for The Cognitive Test

Both of table (4) and figure (3) explain that the value of "T" equal "10.369" for the cognitive test ,and it is a value statistically significant at the level of 0.01 for the post test, where the average of the student's scores of students in the post application is "10.544", while the average of student's scores in the pre application is "1.631" and thereby achieving a second hypothesis.

11.3 The Third Hypothesis

The third hypothesis state as follows:

“There are significant differences between the average scores of students in the skills test pre and post of the educational unit application for the post application ”.

To investigate this hypothesis has been applied test "T" and the following table explains this:

Table 5: Significant of The Differences between The Average Scores of the Students in The Pre and Post Application for The Skills Test

The skill s test	Mean (m)	Std. Devia- tion	N	d f	t	Sig
The first axis: the bodice						
pre	5.801	2.228	15	14	21.882	0.01 for the post
post	59.006	6.111				
The second axis: the bodice						
pre	2.971	0.544	15	14	11.562	0.01 for the post
post	11.129	2.088				
The third axis: the bodice						
pre	3.665	1.565	15	14	13.267	0.01 for the post
post	18.367	1.750				
The fourth axis: the bodice						
pre	4.174	1.441				0.0

post	19.481	3.304	15	14	14.095	1 for the post
The fifth axis: the bodice						
pre	16.611	5.537				0.01 for the post
post	107.983	9.029	15	14	28.613	

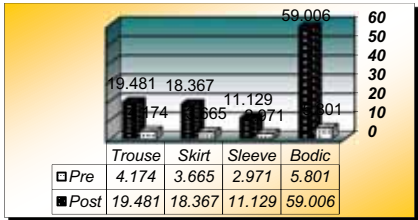


Figure 4: Explain The Differences between the Average of The Student's Scores in The Pre and Post Application for The Skills Test

Table (5) explain that

11.3.1 The value of "T" equal "21.882" for axis I : the bodice ,and it is a value statistically significant at the level of 0.01 for the post test, where the average of the student's scores of students in the post application is "59.006", while the average of student's scores in the pre application is "5.801".

11.3.2 The value of "T" equal "11.562" for axis II : the sleeve ,and it is a value statistically significant at the level of 0.01

for the post test, where the average of the student's scores of students in the post application is "11.129", while the average of student's scores in the pre application is "2.971".

11.3.3 The value of "T" equal "13.267" for axis III : the skirt ,and it is a value statistically significant at the level of 0.01 for the post test, where the average of the student's scores of students in the post application is "59.006", while the average of student's scores in the pre application is "5.801".

11.3.4 The value of "T" equal "14.095" for axis IV : the trouser ,and it is a value statistically significant at the level of 0.01 for the post test, where the average of the student's scores of students in the post application is "19.481", while the average of student's scores in the pre application is "4.174".

11.3.5 The value of "T" equal "28.613" for the total of the skills test, and it is a value statistically significant at the level of 0.01 for the post test, where the average of the student's scores of students in the post application is "107.983", while the average of student's scores in the pre application is "16.611", thereby achieving a third hypothesis.

1. CONCLUSIONS

Fitting is one of the important criteria for consumers in their buying decision, for that the main concern of customers especially teenagers and women is to purchase garments that have proper fitting and yet fashionable which is according to the person's body structure, and clothes are not only for body protection and covering, but also have social and emotional aspects

attached to them, so the authors applied the educational unit on a representative sample of students from the Department of Clothing and Textile, Faculty of Arts and Designs - King Abdul Aziz University, to determine the effectiveness of the educational unit in the acquisition of knowledge and skills for the students and we found the value of "T" equal "28.613" for the total of the skills test, and it is a value statistically significant at the level of 0.01 for the post test, where the average of the student's scores of students in the post application is "107.983", while the average of student's scores in the pre application is "16.611", thereby achieving the hypothesizes of the research.

ACKNOWLEDGMENTS

Technical support:

Hatem Refaie, Professor and Head of the Department of Clothing and Manufacturing, Faculty of Home Economic, Helwan University.

Farag Elmenuawi, Graphic Engineer in Fashion Education technology.

REFERENCES

- [1] Ashdown.S.P and DeLong.M. . Perception testing of apparel ease variation. *Applied Ergonomics* . Vol. 26 No.1:47-54, 1995.
- [2] Barbara A. Frey and Jann Marie Sutton. A Model for Developing Multimedia Learning Projects, *Merlot Journal of Online Learning and Teaching* Vol. 6, No. 2,503: June 2010
- [3] E. Goldsberry, S. Shim and N. Reich. Women 55 years and older. Part II. Clothing and Textiles Research Journal. 14(2), 121-132, 1996.
- [4] Elizabeth G. Liechty, Della N. Pottberg and Judith A. Rasband. fitting & pattern alteration-a multi-method approach . Fairchild fashion & merchandising group, New York, 3rd edition: 99, 1992.
- [5] G. Moore and G Kearsley. Distance Education- A Systems View, 2nd edition: 12, 2004.
- [6] John Barrett, Ph.D. An interview .Oregon, M. Cog. Sc. (UNSW) - THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY:1995.
- [7] Karen Morris. Sewing Lingerie That Fits: Stylish Underwear, Sleepwear, and Loungewear for Everyday Living , Copyright by The Taunton Press:61,2001.
- [8] Kurt Salmon Associates. Annual Consumer Outlook Survey. Atlanta, GA,(1999).
- [9] Kurt Salmon Associates. Annual Consumer Outlook Survey. Paper presented at a meeting of the American Apparel and Footwear Association Apparel Research Committee, Orlando, FL, November ,2000.
- [10]L.J. Anderson (PI), E. L. Brannon, P.V. Ulrich, A.B. Presley , Dave Woronka , M. Grasso and D. Stevenson . Understanding Fitting Preferences of Female Consumers - Development an Expert System to Enhance Accurate Sizing Selection - National Textile Center Annual Report: November 2001.
- [11]Phyllis Brackelsberg and Ruth Marshall. Unit Method of Clothing Construction.waveland press,USA,Illinois ,58:1999.
- [12]Teresa Correia, André Moreira and Isabel Martins. Multimedia resources as a complementary tool of teaching and learning. Case study of a game designed to teach immunology contents in Eunis Conference, 23-25 June, 2010.
- [13]Todd Brashears ,Cindy Akers and James Smith. The Effects of Multimedia Cues on Student Cognition in an Electronically Delivered High School

ABOUT THE AUTHORS

1. Hazem Abdelfattah Abdelmonem
Abdelfattah
Assistant Professor of Clothing Man-
ufacturing
Helwan University.
Faculty of Home Economic.
Department of Clothing and Textiles.
bravehazem@yahoo.com
2. Thanaa Mustafa Aref Alsarhan
Assistant Professor of Clothing Man-
ufacturing
King Abdul-Aziz University.
College of Art & Design.
Department of Fashion Design.
talsarhan@kau.edu.sa

THE USE OF PHOTOGRAPHS IN THE TEACHING/LEARNING OF DESCRIPTIVE GEOMETRY

Bożena KOTARSKA-LEWANDOWSKA

Gdansk University of Technology, Poland

ABSTRACT: The concept of enriching the Descriptive Geometry (DG) course with photographs came from the desire to refresh present learning resources. It also emerged from the conviction that simplified real-life engineering tasks may considerably stimulate students' interest in a given subject. In such an approach the used photographic images used for the exercises should be tightly linked to engineering structures, the given specialization and the surrounding world. At the Civil and Environmental Faculty in Gdansk University of Technology the idea of using more photographs is gradually implemented into the DG course. The paper presents the results in the context of four different majors within the faculty: Civil Engineering, Environmental Engineering, Transportation, Geodesy and Cartography. The differences in the course programs among them allowed to achieve variety of applications in the drawing tasks and projects.

Keywords: descriptive geometry, perspectives, engineering education, photographs.

1. INTRODUCTION

The main task of teaching DG is to develop three dimensional thinking and prepare students for their design work in 3D space. In DG course the issues and topics of tasks should be well balanced (neither too limited and restricted only to standard constructions, nor requiring over-complicated manual construction). It is therefore very important that the course is engaging for students, gives the possibility of their own invention and at the same time pursues educational goals. The topics of the tasks should be regularly changed, so that the students are facing new issues every year.

The contemporary teaching refers to the theory of knowledge by which everyone learns in their own individual way. Therefore, taking into consideration different styles of learning the course should contain the most diverse teaching aids and materials. For DG there are mainly visual resources such as: drawings, real or virtual models, photos, animations and films etc.

The main concern of this paper is the presentation photographs as one of the aspects supporting the teaching/learning process and

simultaneously depicting connections between DG and the real practice.

The article includes the author's own photos, the satellite photographs, the panoramic view of the PGE Arena Stadium in Gdansk, Poland and the images of the Students' Scientific Society of Geometry and Engineering Graphics KRESKA.

2. TEACHING/LEARNING MATERIALS

Teaching materials (teaching aids) are items that provide students specific sensory stimuli that shorten and diversify the learning process, create impressions and insights that allow to transmit more messages in shorter time [2]. Their importance relates to the organization of teaching/learning environment, that targets to achieve the required competencies at a given stage of education.

Learning materials mean the text, visual material, together with a set of copyright exercises prepared by the teacher for specific recipients.

In the DG course photographs are teaching aids which can serve the creation of learning materials and therefore for the purposes of this

publication it is not important to discriminate between teaching and learning materials. They are closely interconnected in the education process wherein the most important are: cognitive, teaching and motivating functions [2]. (Tab. 1)

Table 1: Main functions of the teaching materials.

Function	Meaning	Example
Cognitive	Teaching materials support students' direct knowledge of specific fragments of reality.	Photographs of real architectural and engineering objects, fragments of constructions.
Teaching	Teaching materials are an important source of acquiring the knowledge and skills, facilitate the understanding of educational content, help the learner in making generalizations.	Photographs of models, photographs of drawing constructions, photos of archival drawings, photographs as perspective views, satellite photographs.
Motivating	Teaching materials evoke a positive attitude to learning through inspiring curiosity and interest in the object of knowledge.	All photographs close to student's interest, plans, aspirations, hobbies.

The use of photographs does not only facilitate the learning process, but also supports students' knowledge of the real world, and above all inspires to increase effort. In addition, photographs may be important reminders of activities in the classroom e.g. pictures of real models that are available at a display at the university (Fig. 1).



Figure 1: The photographs of real models, author's photograph.

If it is only possible the photographs of the archival students' drawings from the last century should be available. The majority of them are of the highest craftsmanship although in many cases the thematic scope is far beyond the contemporary program. Such photos can assist well in understanding of some of the construction methods.

3. PHOTOGRAPHS FOR PERSPECTIVES

The photographic images are used not only for visualization of spatial issues, but mainly as a starting point for spatial analysis in perspectives. The photo image as a record of central projection of a real space can be useful for presentation and analysis of the properties of perspective. In the perspective view such elements may be determined: a horizon, a centre point of vision, vanishing and measuring points of principal edges of the object. The topics of tasks in perspective are related to different majors within the faculty.

The Transportation students solve issues referring to visualization of traffic routes. In the sample task the image shows the perspective of two-lane express road. The lane width is 3.7 m, the dash length 4 m and gap between the dashes of the horizontal road markings is 8 m. In the solution the width of median strip (area separating opposing lanes), emergency lane and shoulder of the road should be determined. Additionally, the height of lamps and their spacing should be estimated (Fig. 2).

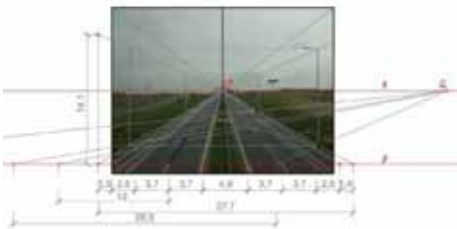


Figure 2: One-point perspective of the road, author's photograph. Task: determining measurements of the elements of the roadway.

The principles of central projection in the application of measuring equipment constitute the theoretical base for the main professional subjects at the faculty of Geodesy and Cartography. Therefore the tasks include not only the construction of a perspective view, but also the reconstruction of an object from a given perspective image.

In the presented task, we assume that the

picture shows a vertical perspective of the small building. The referential points are located on the wall of the building. The ground floor plan and the visible elevations of the building should be determined in the given scale (Fig. 3).

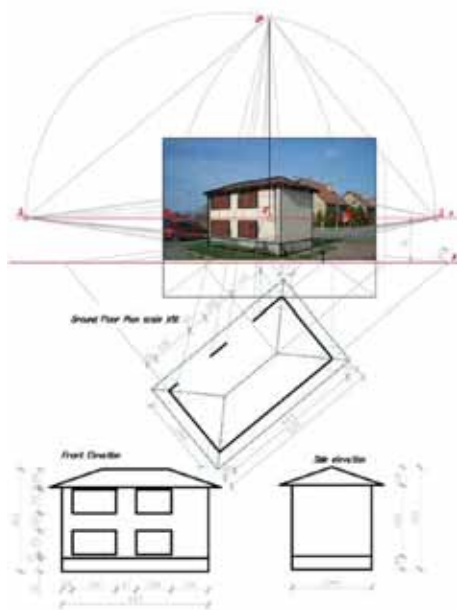


Figure 3: Two-point perspective of the shed, author's photograph. Task: determining measurements of the plan and elevations.

The photographs were taken by the properly mounted camera with centrally built-in lens and for that reason the centre point O^r is located at the geometric center of the original photo at the point of intersection of the diagonals. The photographs that are the basis for the drawing exercise of restitution of a two-point perspective were taken with extreme precision using the professional tripod with the levelling head [3].

For the images in the central projection the accuracy of the photos is of fundamental importance because the image of the object is the part of the drawing. The detailed conditions of

the photos used for restitution of perspective are not listed in the article, but the point is that it should be possible to determine the elements of the reference system and the vanishing and measuring points of the main directions. In practice precise one- or two-point perspective images are quite difficult to obtain.

The restitution results from the perspective image are subject to errors. The size of the error depends inter alia on the accuracy of the photograph, the scale and the precision of drawing.

Nowadays restitution of the perspective views is also becoming increasingly necessary since it can lead to virtual reconstruction as 3D realistic representation of the objects [1, 5, 6].

4. PHOTOGRAPHS FOR PARALLEL PROJECTIONS

In the context of parallel projections photographs may constitute visualization of spatial issues or become an inspiration for drawing tasks such as analysis of the geometric forms of existing facilities and performance of orthogonal or axonometric views of structural elements e.g. fragments of steel structures, connection of two pipes etc. Photographs may concern small items (Fig. 4) however we may expect higher enthusiasm from students towards tasks with more spectacular objects.



Figure 4: The roof ventilation, author's photograph. Possible task: determining orthogonal and axonometric views.

To increase the interest of students the task was located in a newly built football stadium of the PGE Arena in Gdansk, which students may observe from university windows (Fig. 5).



Figure 5: PGE Arena Stadium in Gdansk, Poland, author's photograph.

In the presented example the interior picture of the steel construction of the PGE Arena Stadium is taken from the virtual panorama tour available online [4] (Fig. 6).

The task is to produce a system of cylinders that represents a simplified fragment of the Stadium roof construction (Fig. 7). A correct solution to the task constitutes of an intersection of cylinders and three views in orthographic projection (Fig. 8).



Figure 6: Interior view of the steel construction PGE Arena Stadium in Gdansk, Poland [4].



Figure 7: PGE Arena, the stadium in Gdansk, Poland [4]. Simplified construction as the topic of an exercise, author's own digital model.

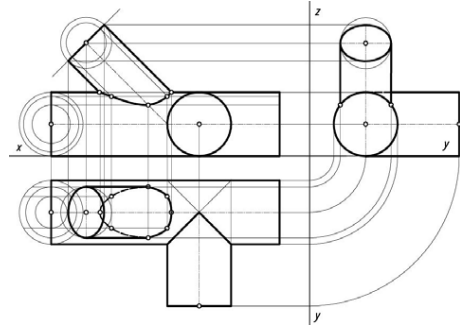


Figure 8: Views of the intersection of pipes in steel construction, orthographic projection. Task: determining the intersection line of three cylinders.

The visualization in oblique axonometric projection is produced with the assumption that the coefficient of reduction $\lambda=1$ for each axis. The points of tangency of ellipses with the contour of cylinders are constructed from the affine transformation of the circle and ellipse (Fig. 9).

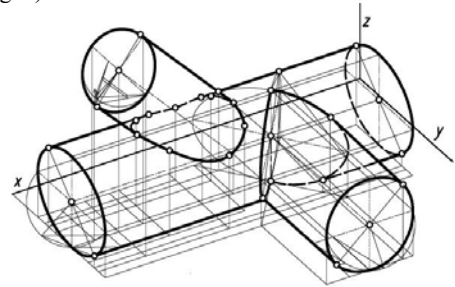


Figure 9: Intersection of pipes in steel construction, axonometric projection. Task: intersection line of three cylinders.

Another example of the use of online resources can be satellite images of existing roofs. Surfing through satellite images creates possibilities for tasks with varying degrees of difficulty, from the simplest to the more complex (Fig. 10). Quick, online access can also be an

inspiration for the students for their own exploration of the subject, especially in the field of the critical analysis of the built roof constructions and the observation of such solutions as gable walls, windows-dormers, changing levels of the base of the roof for different parts of the building (Fig. 11).



Figure 10: The satellite photograph of the simple roof, google maps. Task: The construction of the roof.

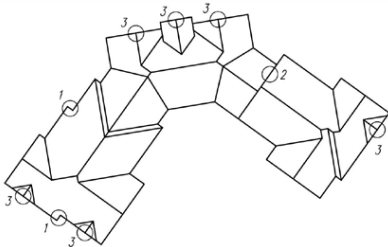


Figure 11: The satellite photograph of the roof, google maps. Task: The construction of the roof. 1-different level of the roof base, 2-gable walls, 3-window-dormers.

Although the satellite pictures are taken from a significant height, the direction of pro-

jection is not perfectly perpendicular to the base. Hence it has to be acknowledged that they represent only the approximate proportions of the roof edges.

5. EXTENDED PROJECTS

Creation of a modern learning environment requires preparation of diverse materials which consider various styles of learning, fields of interest and levels of difficulty. It requires also a set of extracurricular problems in modelling and visualisation. These exercises can be considered by the students in a scientific society. Students who desire to develop their geometric skills using graphical software can pursue independent research and advanced projects in the Students' Scientific Society for Geometry and Engineering Graphics, which for a few years has worked in the Faculty of Civil and Environmental Engineering in Gdansk University of Technology.



Figure 12: The historic campus building.



Figure 13: The model of a historic campus building, Gdansk University of Technology, Group Project of the Students' Scientific Society KRESKA 2011.

One of the group projects within the Students' Scientific Society of Geometry and Engineering Graphics KRESKA was to model a historic campus building in SketchUp and export it to Google Earth database. The photographs of the building were used for textures and added to the virtual model in order to provide a realistic visualisation (Fig. 12, 13).

6. CONCLUSIONS

The concept of using photographic images in the DG course is not recent, but its potential has not been fully explored yet. The presented ideas are an emerging project, still in progress. However it can already be recognised that through implementation of photographic images into the course teachers may find new concepts for their lectures. New tasks inspired by photographs will prevent routine and the DG course is likely to be enriched and refreshed.

In creating extensive learning environment photographs may serve as a bridge between past antecedent drawings and future digital technologies.

They might also stimulate students, who themselves can pick topic, take or look for pictures. Hence photography based exercises may help in successfully developing creativity, independence and problem solving skills.

At the same time those exercises depict applicability of DG, geometric problems as well as enhance observation of geometric forms in the immediate environment. In practical tasks, it is important and necessary to combine everyday experiences with a geometrical analysis and descriptive methods.

REFERENCES

[1] M. E. Abdel-Latif, A. A. Elsonbaty. *Orientation Parameters and Reconstruction of Space Models from a Single Photo*. J. Geometry Graphics 13/1, 91-99 (2009).

[2] F. Bereźnicki. *Dydaktyka kształcenia ogólnego*, Wydawnictwo Impuls, Warszawa 2007.

[3] Z. Brzosko. *Wykreślna restytucja perspektywy*. Wydawnictwo Naukowo-Techniczne, Warszawa 1995.

[4] http://pgarena.gdansk.pl/wirtualny_spacer

[5] H. Stachel. *Descriptive Geometry Meets Computer Vision – The Geometry of Two Images*. J. Geometry Graphics 10/2, 137-153 (2006).

[6] V. Stojaković. R. Štulić. *Virtual Reconstruction of the Kljajićevo Chapel*. J. Geometry Graphics 14/1, 81-91 (2010).

ABOUT THE AUTHORS

1. Bożena Kotarska-Lewandowska, Ph.D., /is an Assistant Professor in the Civil and Environmental Faculty, Gdansk University of Technology. Her research interest are in Descriptive Geometry, CAD, Spatial Visualization and Educational Measurement. She can be reached by mail: bokot@wilis.pg.gda.pl, tel.:+48583471755 or through postal address: Department of Structural Mechanics and Bridge Structures/ Civil and Environmental Engineering Faculty/ Gdansk University of Technology /G. Narutowicza 11/12 / Gdansk 80-233 /Poland.

USING COMPUTER TECHNOLOGIES FOR EDUCATIONAL PROCESS ON GRAPHIC SUBJECTS AT MPEI

Elena KASATKINA, Yuriy STEPANOV and Viktor KAURKIN

Moscow Power Engineering Institute, Russian Federation

ABSTRACT: This paper presents the methods and results of using computer technologies (CT) in educational process on Descriptive Geometry and Engineering Graphics at MPEI. Original methodical instructions, special graphic exercise books (GEB) and individual graphic assignments (IGA) have been developed at the Engineering Graphics (EG) Department of MPEI using computer technologies. We have used such CAD Systems as Autodesk AutoCAD, Autodesk Inventor Series, SolidWorks and others.

Keywords: Instructions, descriptive geometry, engineering graphics, graphic exercise books (GEB)

1. INTRODUCTION

According to Russian Federal Educational Standards the syllabuses of graphic subjects depend on the directions of professional education. The syllabuses for Mechanical Engineering and Power Mechanical Engineering directions are the most complex. So Digital Educational and Methodical Complexes (DEMC) on graphic subjects for these directions are taken into consideration in this paper.

2. POSITIVE AND NEGATIVE EFFECTS OF COMPUTER TECHNOLOGIES

Experience has shown that computer technologies have both positive and negative effects on the results of education. So finding the way to avoid negative effects becomes vital.

2.1 Positive effects

We should consider positive effects of computer technologies in educational process both for students and lecturers [4].

The most important positive effects as lecturers think are the following:

- a very high quality of illustrations for lectures and practical classes;
- the opportunity to use multimedia technologies and demonstrate the

process of analyzing geometry of complex technical objects;

- the opportunity to construct 3D models of complex technical objects and to make reading there shapes easier;
- the opportunity to systematize and to store any methodical and educational matters;
- the opportunity to organize a distance educational process.

2.2 Negative effects

There are such negative effects of using computer and multimedia technologies in teaching process on graphic subjects as:

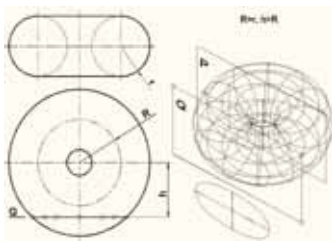
- difference between the rate of expounding information by the lecturer during a lecture and the rate of taking it in by the student;
- students have no motivation to write down the information during the lecture, they can get it from the teacher in a printed or a digital form. This fact limits modes of perception information during the lecture;
- lecturers have faced the problem of passive contemplating the lecture by students. We have called it

"cinema effect".

Understanding of importance of developing motivation, attention, long- and short-term memory and space imagination have led us to special methods to avoid the above mentioned negative effects.

To synchronize process of transmitting and perceiving information during lectures and practical classes special students GEB are used [1, 2] and [5]. Such books are published for every stage of graphic education. GEB includes base terms, definitions, graphical assignments, 3D models for lectures, practical classes and self study.

One of the pages from GEB on the subject of "Lines of Intersection of a Surface and a Plane" is shown in Figure 1.



Blank reserved for writing notions and their definitions, develop algorithms for solving problems

Figure 1: The page with assignment from GEB on descriptive geometry.

During a lecture the given problem is solving in step with the lecturer and blanks are filling in. The page of GEB after solving the problem is shown in Figure 2.

The Digital Educational and Methodical Complex (DEMC) on descriptive geometry made at our University includes such modules as GEB, digital manual, block of tests, individual and final exam assignments. This DEMC is described in [6]. We use different

variants of assignments both for self studying, in-course assessment and final examination.

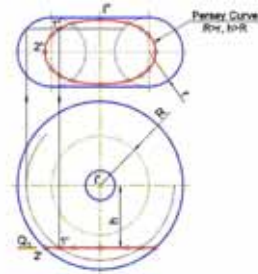


Figure 2: The problem solved in GEB.

Lecturer usually use both frame and realistic 3D models to explain a theory as it is shown in Figure 3.

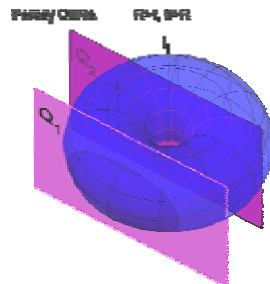



Figure 3: Realistic 3D model.

3. USING MULTIMEDIA TECHNOLOGIES IN LECTURES ON ENGINEERING GRAPHICS

On the next stage of engineering education our students study the subject of Engineering Graphics. They construct 2D and 3D models of real technical objects, develop such design documents as: sketches of details, theoretical drawings of units, specifications, assembly drawings, detail drawings. Special digital reference book, developed at Engineering Graphics (EG) Department of MPEI is used on

So on this stage of education multimedia technologies are used in lectures on Engineering Graphics. In number of cases it considerably simplifies given theoretical material. It may be illustrated on the example of learning the subject "Assembly drawing and Specification".



When working with 2D drawing of the assembly unit in GEB and learning its image on projection screen under the lecturer's guide students acquaint with assembly unit construction and working principle. During the lecture special attention is paying to studying simplifications in assembly drawings and to

The digital data base developed at MPEI is used to select standard products included into the assembly units. Any representations from the data base can be used and edited for including into the assembly drawing.

8000

Technical drawing of the 8000 series ball valve. The drawing includes three views: a front view (top left), a side view (top right), and a top view (bottom). Dimensions are provided in millimeters (mm) and inches (in).

- Front View:** Shows the internal components of the valve, including the ball, stem, and handle. Dimensions include overall height (A), handle length (B), and various internal diameters and lengths.
- Side View:** Shows the profile of the valve body. Dimensions include overall width (C) and height (D).
- Top View:** Shows the circular face of the valve body. Dimensions include outer diameter (E), inner diameter (F), and distance between mounting holes (G).

DIMENSIONI E TOLLERANZE

Dimensione	Tolleranza
A	+0.10/-0.05
B	+0.10/-0.05
C	+0.10/-0.05
D	+0.10/-0.05
E	+0.10/-0.05
F	+0.10/-0.05
G	+0.10/-0.05

* Valore per coppia

4. ASSESSMENT

The assignment for the final examination on descriptive geometry is in Figure 6. It can be fulfilled by the student using traditional or computer as he wishes. He can use 2D and 3D models to solve the problem but both models must be presented.

1318

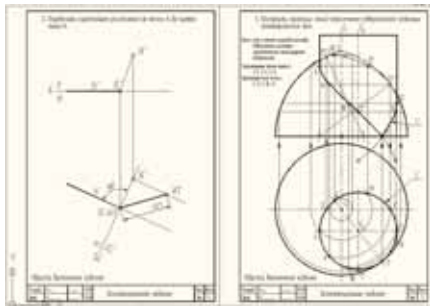


Figure 6: The final examination assignment (fragment).

5. CONCLUSIONS

Experience has shown that computer technologies have both positive and negative effects on the results of education. Special methods to avoid the negative effects are used at MPEI.

Special graphic exercise books (GEB) help to synchronize transmission and perception of the information and leads to the deeper comprehension of the lecture.

Digital Educational and Methodical Complexes (DEMC) including GEB, assignments for self-study, testing, examining are developed and successfully used at MPEI.

It was reached the balance between traditional and modern information technologies in educational process on graphic subjects.

ACKNOWLEDGMENTS

The authors acknowledge T. A. Bobrova and I. V. Gordeeva for great contribution to the paper.

REFERENCES

[1] T. A. Bobrova, N. A. Burdunina, A. O. Gornov and others. Graphic exercise book for lectures and practical classes on the subject of "Base of developing design documentation". – Moscow, Publishing house of MPEI, 2010 – 54 p.

[2] T. A. Bobrova, L. G. Golovina, I. V. Gordeeva and others. Graphic exercise book for lectures and practical classes on the subject of "Theory of Draw Constructing". – Moscow, Publishing house of MPEI, 2013 – 80 p.

[3] T. A. Bobrova, E. P. Kasatkina. Using AutoCAD System opportunities for illustrating information on lectures and practical classes. *Proceedings of the 19-th International Science and Technical Conference "Information Means and Technologies". Moscow, Russian Federation, October 18-19*, pages 112–117. MPEI University publishing house, 2014.

[4] S. A. Serkov, E. P. Kasatkina. Information technologies in bachelor and master of science educational process for Power Machine Engineering direction. *Proceedings of the International Science and Methodical Conference on informatization of engineering education "INFORINO 2014". (Moscow, Russian Federation, April 15-16)*, pages 151–154. MPEI University publishing house, 2014.

[5] Y. V. Stepanov. Graphic exercise book for lectures and practical classes on the subject of descriptive geometry. – Moscow, Publishing house of MPEI, 2011 – 80 p.

[6] A. I. Yevseyev, B. R. Lipai and S. I. Maslov. Informatization of engineering education: digital resources of MPEI. – Moscow, Publishing house of MPEI, 2009 – 190 p.

ABOUT THE AUTHORS

1. Elena Kasatkina - the Head of the Engineering Graphic Department of the National Research University "Moscow Power Engineering Institute" (MPEI).
2. Yuriy Stepanov - associated professor of the Engineering Graphic Department of MPEI.
3. Victor Kaurkin - PHD, associated professor of the Engineering Graphic Department of MPEI.

USING OPTICAL ILLUSION PATTERNS AFFIXED TO TOY BLOCKS FOR LEARNING HUMAN ERRORS IN THREE-DIMENSIONAL PROJECTION

Tomoko OHTANI¹ and Kazushi MARUYA²

¹ Research Institute of Electrical Communication, Tohoku University, Japan

² Communication Science Laboratories, Nippon Telegraph and Telephone Corporation, Japan

ABSTRACT: We propose a workshop program for learning discrepancies between the 3D object's image as envisaged using a 2D image and the actual finished object. Participants of this workshop use an optical illusion block, which is a set of interlocking blocks textured with geometric patterns that induce optical illusions. Participants assemble the optical illusion blocks into solid objects (illusion objects) and evaluate them. In addition, they create objects that have the same shape as their illusion objects using non-patterned blocks (control objects) and then evaluate them. We conducted a series of workshops and confirmed that participants can make objects that cause optical illusions by viewing them from a specific angle. This was shown in the significant difference between evaluation scores for illusion and control objects. These results suggest that participants in the proposed workshop program repeatedly performed the conversion between the 2D-based mental image and physical 3D objects. In addition, the results suggest that participants can learn how to utilize discrepancies between 2D and 3D images.

Keywords: optical illusion, construction toy, structure, spatial perception, graphic science education.

1. INTRODUCTION

In three-dimensional (3D) object design, when non-experts model a 3D object based on the object's drawing, the resulting object occasionally differs from the designer's intention. To avoid such a situation, the designer is required to recognize the possibility of human error in the mental conversion of 3D objects into a 2D image and have the skill to correctly map each point of view to the 2D plan of the imagined 3D object. Moreover, object designs that cause an optical illusion for the observer (e.g., Varini's architecture [1]) have been recently incorporated. In drafting such objects, the designer needs more detailed knowledge and skill to perform the 3D conversion, considering the observer's viewpoint and characteristics of human visual processing. Therefore, teaching technical skills for mental 3D conversion along

with the characteristics of human visual processing is quite important in graphics education [2]. However, teaching methods supporting the acquisition of such knowledge and skills have not been established in this field. To the best of our knowledge, neither a systematic method to quantitatively evaluate student behavior nor a method that incorporates characteristics of visual processing has been proposed thus far.

We propose a workshop-based educational program for learning discrepancies between the 3D object's image as envisaged using a 2D image and the actual finished object, making use of toy blocks with optical illusion patterns affixed to their surfaces, called "optical illusion blocks" [3; Figure 1]. In this program, the participant first creates an object that uses the optical illusion blocks (an illusion object) and, using the same configuration, an object that uses the non-patterned blocks (a control object).

While making the object, they view the object from various angles. This detailed observation enables participants to notice that the same 3D object can appear differently because of changes in pattern combinations in the 2D retinal image. Further, they observe that the 3D object's physical structure and mental image based on the 2D view can be inconsistent. Adding to these making activities, participants are asked to conduct a subjective evaluation of the created objects using a semantic differential method [4, 5]. The obtained data can be used in subsequent quantitative analysis of the created object attributes.



Figure 1: Optical illusion blocks.

2. OPTICAL ILLUSION BLOCKS

We used two types of toy blocks: One is a 1 x 6 stud rectangular block. The other is 2 x 2 stud flat rectangular block.

We affixed to the surface of the toy blocks three optical illusion patterns, which were the Café Wall illusion, Shape from shading, and illusory contour, to the surface of toy blocks.

Café Wall illusion (see Figure 2) is an illusion of perceived orientation that is induced by displaced rows of black and white squares. The gray lines between the rows are physically parallel. However, they are perceived to be tilted [6, 7]. This illusion also occurs when the rows

consist of sinusoidal gratings, instead of square gratings consisting of black and white squares.

We affixed a gray line and a single row to each sides of the 1x6 stud block (Café Wall block). A gray line was aligned to the upper horizontal long edge of the block. The position (spatial phase) of square rows were displaced by half the width of square (1/2 unit). When two blocks are combined to make the two rows flanking with a different phase, the illusion occurs and the shape of the toy block, as well as the orientation of the toy block, appears to be skewed.

Shape from shading (SFS) patterns [8, 9] are shown in Figure 3a. These patterns consist of rectangle or diamond figures on a uniform gray background. Two edges of these figures are drawn in white, and the other two edges are drawn in black. These configurations of black and white edges cause an optical illusion of depth. The rectangle and diamond figures surrounded by these edges (SFS figures) appear to pop out of or sink into the background surface. The direction of depth illusion depends on the direction of the SFS figure. For example, when the white two edges of the rectangle are in the bottom and the left, the SFS figure tends to appear concave. When the two white edges are located in the top and the right, the SFS figure tends to appear convex.

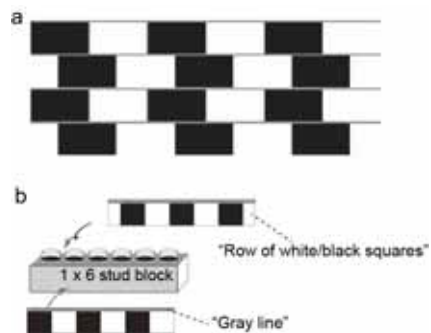


Figure 2: (a) The Café Wall illusion and (b) the Café Wall block.

We affixed the SFS figures to each side of a cube consisting of two 2x2 stud blocks (Figure 3b; SFS block). It is known that the depth illusion is not independent for each of the figures; however, is determined consistently across a large area of the visual field. Additionally, the depth illusion interacts with external information on the light source direction [5, 6]. As the number of elements in the same direction increases, the certainty in the estimation of the light source direction will increase. In fact, this depth illusion is stronger and more stable when we see the multiple figures sharing the same direction than when we see only one SFS figure.

Furthermore, aiming to employ another type of optical illusion, we prepared different sizes of SFS figure. When the observers see the object, the image size projected on the retina becomes larger as the viewing distance is decreased.

Our visual system uses this physical relationship in estimating the distance between the eye and the object. When we see two figures have a same texture but a different size, we perceive them as if they are displayed at different depths. The objects consisting of SFS blocks with different size of SFS figure could cause this other type of depth illusion.

An illusory contour pattern is shown in Figure 4a. The elements in the pattern consist of an oriented white line drawn on a black circle. When these elements are appropriately located, the white lines are perceptually connected, and illusory white lines are perceived on the white background.

We affixed an element to each side of a cube consisting of two 2x2 stud blocks (“i-Contour block”). By composing the i-Contour block in a certain way, the white lines can be perceptually connected (Figure 4b). As shown in the figure 4b, the formation of illusory contour is not limited within the same physical surface but can occur over the line segments at different depths. Such a formation of an illusory plane constructed from perceptually connected lines is

sometimes utilized in designs of buildings or constructions (e.g., Varini [1]). The i-Contour block can be used to design a simple object that causes this illusory plane perception.

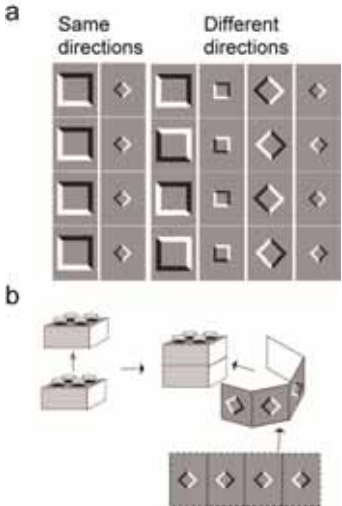


Figure 3: (a) Shape from shading illusion and (b) the SFS block.

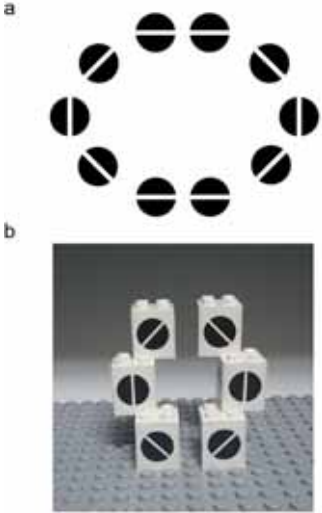


Figure 4: The illusory contour illusion and (b) an example of the object consisted of i-Contour blocks.

3. WHORKSHOP DESIGN

3.1 Workshop protocol

The workshop is held in a quiet room so that participants can easily talk with each other. The protocol consists of four stages (see Figure 5). In advance of beginning to make the object, we briefly lectured the students about an optical illusion so that each participant could utilize the optical illusion effectively in the design of their objects. This purpose was explicitly expressed by the facilitator during that lecture. In the first stage, participants made his/her object using optical illusion blocks (illusion object). The participants were also asked to give a title to their illusion object. In the second stage, they made an object that had the same structure as their illusion object by using only white blocks (control object). After that, they took photographs of their objects from their preferred angle so as to make the object appear to reflect the intention of its creator. In the third stage, they evaluated these objects by using the Semantic Differential method (SD method) [4, 5]. In the fourth stage, each participant made a presentation about his/her object in front of other participants, of duration no longer than one minute. The other participants observed the objects from the viewpoint that each creator recommended. Finally, they had a discussion about their objects.

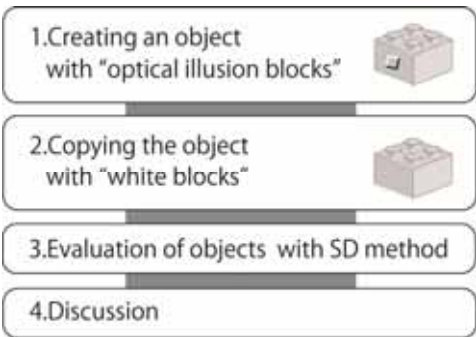


Figure 5: Main workshop flow.

3.2 Evaluations

Participants were asked rate their objects using SD method, which required them to use a number of scales ranging from 1 to 5 with anchoring adjective at each extreme. We used 12 adjective-pairs: soft–hard (hardness), hot–cold (temperature), strong–weak (strength), bright–dark (brightness), honest–dishonest (felicity), pretty–not pretty (attractiveness), large–small (size), ornate–plain (display), familiar–unfamiliar (familiarity), light–heavy (weight), calm–excitable (calmness), disfavored–favorite (favorability). This provided a 2 x 12 factorial design for data analysis consisting of 2 Blocks (the illusion condition / control condition) crossed with 12 adjective-pairs (SD evaluation) design.

4. EXAMPLE WORKSHOP PROCEDURES

To confirm the effectiveness of the proposed workshop, we conducted three workshops that followed the proposed protocol; however, each workshop slightly changed. The protocol used in the first and second sessions and that in the third session were different. These details of protocols are as following.

The first and second sessions were held at University of Tokyo and Tohoku University, respectively. Participants were five males and four females aged 22–41. There were eight trials. The workshop sessions lasted 120 minutes, including a 15–20 minute break time in the middle. In this workshop, participants were asked to repeat the first three stages eight times. In the first four runs, participants used only the SFS blocks to make illusion objects. The number of SFS blocks was five in the first run, and increased by five blocks for each run. Therefore, the participants used 20 SFS blocks in the fourth run. After that, for the latter four runs, participants used both SFS and Café Wall blocks. As in the first four runs, the number of blocks was limited to 10 for SFS and Café Wall blocks in the 5th run, and then increased as 10

SFS-20 Café Wall blocks (sixth), 15 SFS-10 Café Wall blocks (seventh), 15 SFS-20 Café Wall blocks (eighth), respectively.

The third workshop session was held at Miyagi Prefectural Sendai Daiichi High School. Ten high school students, who did not know about the purpose of the research, participated in the workshop. There were nine males and one female, all aged between 15 and 18 years.

After a brief (~10 min) lecture about optical illusions in advance, participants were separated into groups and sat facing each other. There were the optical illusion blocks to use in this investigation provided on the tabletop for each group. In this workshop, each participant was supplied with thirty-four optical illusion blocks as a total. This set consisted of 16 SFS blocks [4-types (big/small and rectangle/diamond) x 4], 9 Café Wall blocks [6 square-wave blocks and 3 sinusoidal-wave blocks], and 8 i-Contour blocks. The participants were allowed to make their objects using only a subset of these blocks. Also, they were allowed to borrow extra blocks from other participants.

In this workshop, participants ran each stage of the protocol only once. At first, participants made their construction using optical illusion blocks within 20 minutes. Next, they gave their constructed object a title. Then, within another 10 minutes, participants made an object with the same structure as their illusion object, but using only white blocks. After they took photographs, they evaluated these objects using SD method, taking approximately 15 more minutes. When all participants had finished this process, another 45 minutes were taken up as they each made presentations and held discussion. As a whole, the workshop took 105 minutes including the 15-minute break in the middle.

5. RESULTS AND DISCUSSION

5.1 Examples of constructed illusion objects

Figure 6 shows examples of the illusionary objects made by participants in the workshop series.

The typical usage of SFS blocks is shown in Figures 6a and, 6b. The example object having bumpy forms consisted of SFS blocks arranged in alternating directions. The bumpy impression is enhanced by the SFS pattern arranged in this way. Figure 6e shows another example in which two SFS blocks, in the second rows from the bottom, are arranged in the same position. However, it seems that two depths are different according to the direction of the optical illusion pattern.

The Café Wall blocks were used to make a object that are perceived to contain slant structure (see Figure 6c, 6d, 6e). In Figure 6c and 6e, the participant presented slope rails with the Café Wall blocks. In Figure 6d, the Café Wall blocks were used to make walls in a slope.

The i-Contour blocks were utilized to make subjective contours that sometimes cause perceptual completion. Some participants made illusionary plane surfaces surrounded by those subjective contours (Figure 6c and 6e). Note that a triangular contour appears only when viewing the illusion object from a particular viewing angle. However, there was no the triangular contour in the left side of Figure 7b.

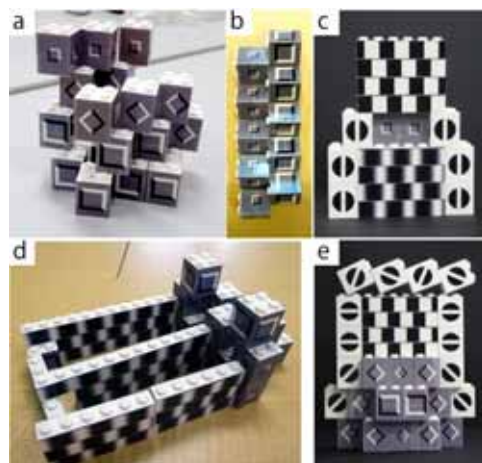


Figure 6: Examples of workshop illusion objects.

5.2 Evaluation of objects with optical illusion blocks

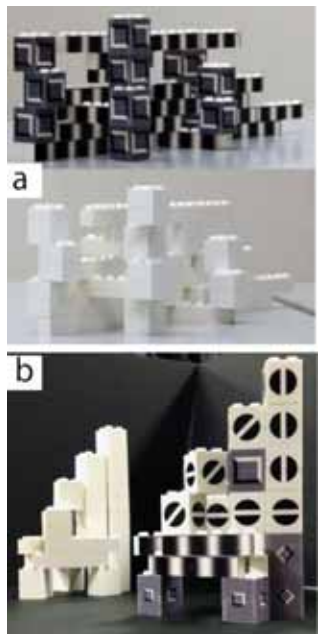


Figure 7: Comparison of an illusion object and a white object.

Figure 7 shows two sets of the illusion and control objects that have the same form. The two object types give different impressions depending upon whether the surface is white or displaying optical illusion patterns, even though the objects have the same form. We examined whether this difference of impression is found quantitatively in the evaluation data.

We obtained a set of evaluation values for 73 objects from the workshop series (at the University of Tokyo and Tohoku University; 63 objects, at Miyagi Prefectural Sendai Daiichi High School; 10 objects). We pooled the evaluation scores for each participant and object. We then converted the raw scores into z-scores for normalization. Figure 9 shows the mean of the obtained z-score for each scale and object type. The mean evaluation scores were significantly

different between illusion and control objects. An ANOVA with the factors of Block and SD evaluation conditions revealed significant main effect for SD evaluation ($F(11, 1991)= 8.99$, $p < .01$) and interaction($F(11, 1991)= 20.63$, $p < .01$); however, no main effect for Block. *Bonferroni* post hoc analysis confirmed that objects in the illusion condition, in comparison to the control condition, gave evaluations that were hotter, stronger darker, larger, more ornate, more familiar, heavier, and more excitable ($p < .01$). On the other hand, the control condition was softer than the illusion condition ($p < .05$). These results indicate that the difference of impression between control and illusion object is well reflected in the evaluation scores.

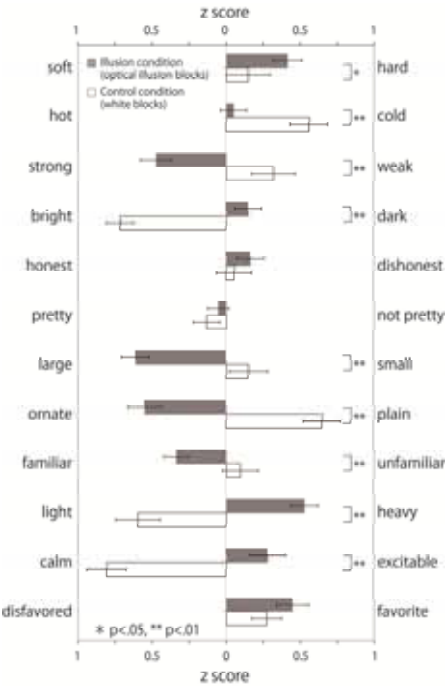


Figure 9: Mean z-score for each SD value. Error bars denote SE; gray bars denote illusion conditions; white bars denote the control conditions.

6. GENERAL DISCUSSION

In this paper, we proposed a workshop plan to learn to recognize the possibility of human error in the mental conversion of 3D objects into a 2D image and to learn the skill to correctly map each point of view on the 2D plan of an imagined 3D object. To examine the efficacy of the proposed plan, we held a series of workshops for adults and high school students. We collected 73 example illusion objects created with optical illusion blocks. In addition, we collected 73 objects with the same form as the illusion objects, but were created with non-textured white blocks. These experimental workshops demonstrated that the series of workshop plans are effective for a wide range of individuals, including high school students. Interestingly, the quality of the students' objects was similar to that of the objects made by the adults. This indicates that the plan of these workshops does not need prior knowledge. The high school participants likely had never heard a lecture about optical illusions, because students in Japan typically first learn about optical illusions during university psychology studies. These high school students had only attended the brief lecture about optical illusions in advance to the workshop.

Our participants in this study were older than 16 and probably mature enough to understand the theoretical aspects of optical illusions. In this regard, the question remains of how early the proposed workshop might be effective in the course of development remains to be addressed.

While participants made trial and error constructions in only a short time, they were able to make creative objects with optical illusion blocks, and to present their intentions and their preferred points of observation. This suggests that the proposed workshop programs are effective in aiding students to learn to recognize the possibility of human error in mental conversion of 3D objects into 2D images and to have the skill to correctly map each point of view on the 2D plan with the imagined 3D object.

The SD method showed that mean evaluation scores for illusion object were systematically different from those for control objects.

This difference not only supports the effectiveness of the present workshop plan, but suggests a systematic relationship between multi-dimensional evaluation scores and combinations of block type, direction and viewing angle. In fact, preliminary findings had suggested the presence of such systematic relationships. For example, an object containing a combination of SFS blocks with a different texture size and direction tended to be evaluated as heavier and larger than an object containing only SFS blocks with the same texture size and combination direction. Additionally, in a combination of other blocks, we found a systematic shift of evaluation scores among the different combinations. Future investigations with larger samples and more a systematic experimental design are required to clarify the nature of these interesting interactions.

7. CONCLUSIONS

We proposed a workshop-based educational program for teaching discrepancies between the 3D object's image as envisaged using a 2D image and the actual finished object using optical illusion blocks. In this paper, we investigated tendencies found in the application of each visual pattern of the optical illusion blocks to the final designs assembled by the participants. The results suggested that participants in the proposed workshop program repeatedly performed the conversion between the 2D-based mental image and physical 3D objects and therefore learned a way to utilize, in this context, the discrepancies between 2D and 3D images.

ACKNOWLEDGMENTS

This work was supported in part by "Hayao Nakayama Foundation for Science & Technology and Culture" and "Tohoku Leading Women's Jump up Project".

REFERENCES

- [1] Felice Varini. <http://www.varini.org/> (accessed 2014-6-11).
- [2] A. Chalmers and C. Dalton. Visual perception in computer graphics education. *Euro graphics-ACM SIGGRAPH Workshop on Computer Graphics Education*, 79-82, 2002.
- [3] T. Ohtani, J. Watanabe and K. Maruya. Three-dimensional design utilizing toy blocks with optical illusion patterns affixed to their surfaces: an interaction between structure and texture patterns on three-dimensional object recognition. *Cognitive Studies* (in Japanese), **17**(3): 580-588, 2010.
- [4] C. E. Osgood, G. Suci and P. Tannenbaum. *The Measurement of Meaning*. Urbana, IL: University of Illinois Press, 1957.
- [5] C. E. Osgood. The nature and measurement of meaning. *Psychological Bulletin*, **49**(3): 197-237, 1952.
- [6] R. L. Gregory and P. Heard. Border locking and the Café Wall illusion. *Perception*, **8**: 365-380 1979.
- [7] Y. Tani and K. Maruya, Reversed Café Wall illusion with missing fundamental gratings. *Vision Research*, **46**: 3782-3785 2006.
- [8] V. S. Ramachandran. Perception of shape from shading. *Nature*, **331**: 163-166, 1988a.
- [9] V. S. Ramachandran. Perceiving shape from shading. *Scientific American*, **259**: 76-83 1988b.

ABOUT THE AUTHORS

1. Tomoko Ohtani is an Assistant Professor at the Research Institute of Electrical Communication, Tohoku University, Sendai, Miyagi, Japan. She received M.A. in 2001 and Ph.D. degree in 2008 from the University of the Sacred Heart in Psychology. Her current research in-

terests include multisensory information processing and computer-human interaction.

2. Kazushi Maruya is a Senior Research Scientist with the Sensory Resonance Research Group, Human Information Science Laboratory, NTT Communication Science Laboratories, Atsugi, Kanagawa, Japan. He received M.A. and Ph.D. degree from the University of Tokyo in Psychology in 2001 and 2005. His current research interests include visual motion processing for perception, especially for perception of dynamic natural scenes. He is also interested in human-computer interface design for reading digital text.

VISUALISING VARIABILITY: CONFIDENCE REGIONS IN LEVEL SET ESTIMATION

Hanna JANKOWSKI¹ and Larissa STANBERRY²

¹York University, Canada

²Seattle Children's Research Institute, U.S.A.

ABSTRACT: Level sets of a function appear in numerous scientific problems, for example, in the detection of areas of high cancer rates. In practice, the true function is unknown and is therefore estimated. Here, we focus on quantifying the risk of replacing the unknown (true) function with its known estimator in the level set. We describe the variability, or accuracy, of the resulting estimator via the statistical notion of a confidence region, which naturally provides a graphical representation of variation easily visualized by the practitioner.

Keywords: Geometry, graphics, statistical inference, level set

1. INTRODUCTION

A level set of a real-valued function f is a set of points where the values $f(x)$ satisfy some constraint. For example, if $f(x)$ is the intensity of radiation at a point x on a surface, then the level set could be the area of the surface where the radiation intensity exceeds some threshold. Level sets are used in a variety of problems, including “hot spot” detection (e.g. to identify regions with low vegetation growth or areas with high cancer rates) and forecasting of extreme weather events.

In practice, the true function f is often unknown and is estimated from observed data by \hat{f}_n . The level set of f is then estimated by the level set of \hat{f}_n . A key practical question is to characterize and quantify the accuracy/risk of using the estimate \hat{f}_n in the level set vs. the unknown true function f . In statistics, this type of error is usually quantified using variance or even more appropriately, using standard deviation. However, characterizing, computing, and visualizing the variability of the level set estimate is a particularly difficult statistical problem. This is, in large part, due to the fact that the space of (closed) sets is nonlinear.

Mathematically, we denote the level set as

$$F(c_1, c_2) = \{x \in \mathcal{D} : c_1 \leq f(x) \leq c_2\}, \quad (1.1)$$

where $\mathcal{D} \subset \mathbb{R}^d$, f is a continuous function, $f : \mathcal{D} \mapsto \mathbb{R}$, and $-\infty \leq c_1 \leq c_2 \leq \infty$. The true level set $F(c_1, c_2)$ is then estimated as

$$\hat{F}_n(c_1, c_2) = \{x \in \mathcal{D} : c_1 \leq \hat{f}_n(x) \leq c_2\}. \quad (1.2)$$

Sets of this form appear in various statistical problems, such as estimation of contour clusters [18, 19] or the estimation of density support [10]. In Section 5, we show how to apply the proposed method to the estimation of the domain of covariates with specified response level(s). Such a situation arises often in medical studies, such as dose optimization and toxic dose estimation [1, 23, 24], as well as in other fields. Additional examples include mode estimation, estimation of highest density and/or intensity regions [17], and abnormal system behaviour detection. The example given in [4] estimates the spherical density of double stars, and uses level sets to find directions with high densities of these double stars. [7] consider highest density regions in the estimation of the wintering location of the wood thrush songbird. Level sets are also closely related to random closed sets, and this relationship was studied in image inference applications in [8, 9].

Here, we propose to use confidence regions to characterize and visualize the variability of the

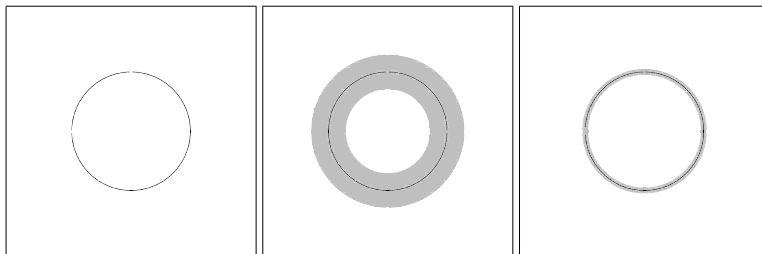


Figure 1: A toy example showing the observed level set (left) and two different confidence regions (centre and right, the confidence region is shown in grey and the estimated level set in black).

level-set estimators. In statistics, a $100(1 - \alpha)\%$ confidence region \mathbf{R} of the true level set $F(c_1, c_2)$ is a random set that covers $F(c_1, c_2)$ with a probability of at least $1 - \alpha$, that is, $\text{pr}(\mathbf{R} \supset F(c_1, c_2)) \geq 1 - \alpha$, under repeated experimentation. In practice, the common choice is $\alpha = 5\%$ that yields a 95% confidence region. A classical example is the estimation of the population mean μ . Under random sampling, the estimate of μ is the sample mean, \bar{X}_n . Estimate \bar{X}_n is the point estimate and it gives, statistically, the best guess of μ based on the observed data. The corresponding 95% confidence interval $\mathbf{R} = (\bar{X}_n - 1.96\sigma/\sqrt{n}, \bar{X}_n + 1.96\sigma/\sqrt{n})$ reflects the accuracy of the estimator \bar{X}_n (σ is a standard error of the sample), i.e. a narrower confidence interval indicates a higher accuracy. The coverage of \mathbf{R} is the probability of $F(c_1, c_2) \subset \mathbf{R}$. In the above example, the coverage of \mathbf{R} is approximately 95%.

Figure 1 shows a toy example of a level set $\hat{F}_n(0, 0)$ (left), along with two different 95% confidence regions (middle and right). In practice, the true level set will not be known, and therefore we do not add it to this figure. The two confidence regions greatly differ in size with the larger one (middle) showing the estimator to be less accurate than the smaller one (right). Clearly, it is important for the practitioner to be able to visualize such a difference in the reliability of the estimate.

In set estimation, confidence regions can be used to describe the accuracy of the set estimator as well as its global and local variability. In [8], we proposed a new approach for calculating confidence regions for the mean of a random set. Developed independently, the approach was closely related to that of [3] as both were based on the idea of supremum inversion. This work is motivated by applications to image inference and effective dose estimation [3, 6, 8, 9, 12, 16]. Here, we provide a general framework for calculation of confidence regions for level sets. We focus on the case where the rate of convergence of the estimated function is \sqrt{n} , which is most often seen in parametric statistics. For some results in the nonparametric setting, we refer to [13].

The outline of this paper is as follows. In Section 1.1 we outline our notation and key assumptions. In Section 2 we consider consistency of the plug-in estimators and in Section 3 we describe our general approach for confidence region calculation under the assumptions of Section 1.1. Section 4 describes a simulation study, and Section 5 provides some examples. Code for the examples is available at www.math.yorku.ca/~hkj/.

1.1 Notation and Assumptions

Unless otherwise stated, we assume that \mathcal{D} is the working domain and write, for example, $F(c_1, c_2) = \{x : c_1 \leq f(x) \leq c_2\}$ without stating

that $x \in \mathcal{D}$ explicitly. We assume that $\mathcal{D} \subset \mathbb{R}^d$, and denote the Euclidean norm of x as $|x|$.

We write $B_r(x_0) = \{x : |x - x_0| \leq r\}$ for the closed ball of radius r centred at x_0 . For a set A , we write A°, \bar{A}, A^c and ∂A to denote its interior, closure, complement and boundary. Unless noted otherwise, set operations are calculated relative to the domain \mathcal{D} . That is, $A^c = \mathcal{D} \setminus A$, and so forth. Furthermore, for a set A , we define $A^\delta = \{x : B_\delta(x) \cap A \neq \emptyset\} = \bigcup_{x \in A} B_\delta(x)$. Deterministic sets are denoted using capital letters A, B, \dots , while bold upper-case lettering, $\mathbf{A}, \mathbf{B}, \dots$, is used for random sets. We do this to emphasize the difference between the random and observed set. Recall also that the Hausdorff distance between two sets, A and B , is defined as

$$\rho(A, B) = \inf \left\{ \delta > 0 : A \subset B^\delta, B \subset A^\delta \right\}.$$

The notation $C(\mathcal{D})$ is used to denote the space of continuous functions $C(\mathcal{D}) = \{f : \mathcal{D} \mapsto \mathbb{R}, f \text{ continuous}\}$ endowed with the uniform metric. We write $X_n \Rightarrow X$ to say that X_n converges weakly to X . When handling weak convergence of stochastic processes or random fields, we assume that they take values in $C(\mathcal{D})$.

Suppose now that \hat{f}_n is a random, continuous function such that

$$(A1) \quad \sup_{x \in K} |\hat{f}_n(x) - f(x)| \rightarrow 0$$

almost surely (almost everywhere), as $n \rightarrow \infty$, for all compact sets $K \subset \mathcal{D}$. To construct confidence regions, the sets (1.1) are restricted to a compact window $\mathcal{W} \subseteq D$, and we require the assumption of weak convergence

$$(A2) \quad \sqrt{n} \{\hat{f}_n(\cdot) - f(\cdot)\} \Rightarrow \mathbb{Z}(\cdot),$$

where $\mathbb{Z}(\cdot)$ is a continuous random field on \mathcal{W} . In practice, assumption (A2) can be checked using the techniques described in [2, 25] for $\mathcal{D} \subset \mathbb{R}$ or [11] for $\mathcal{D} \subset \mathbb{R}^d$.

2. CONSISTENCY

An estimator is said to be consistent if it approaches the quantity it is estimating as the sample size increases. Consistency is a core concept

in statistics, because if an estimator is biased, this bias becomes negligible for a sufficiently large sample size. Below we provide conditions required for consistency of $\hat{\mathbf{F}}_n(c_1, c_2)$. The proofs appear in the Appendix, and follow from [16] and/or [4].

Let \mathcal{F} be the family of closed sets of \mathbb{R}^d and let \mathcal{K} denote the family of all compact subsets of \mathbb{R}^d . For a probability triple (Ω, \mathcal{A}, P) , a random closed set is the mapping $\mathbf{A} : \Omega \mapsto \mathcal{F}$ such that for every compact set $K \in \mathcal{K}$

$$\{\omega : \mathbf{A}(\omega) \cap K \neq \emptyset\} \in \mathcal{A},$$

(cf. [15]). Note that

$$\begin{aligned} & \{\hat{\mathbf{F}}_n(c_1, c_2) \cap K \neq \emptyset\} \\ &= \left\{ \inf_{x \in K} \left| \hat{f}_n(x) - \frac{c_1 + c_2}{2} \right| \leq \frac{c_2 - c_1}{2} \right\}, \end{aligned}$$

Therefore, since the functions \hat{f}_n are continuous almost surely, the estimators (1.2) satisfy the measurability requirement and are well-defined.

A random closed set \mathbf{A}_n converges strongly to a deterministic set A if for any compact set K , $\rho(\mathbf{A}_n \cap K, A \cap K) \rightarrow 0$ almost surely (almost everywhere) [16]. The key conditions for the consistency of the estimators (1.2) are

$$\{x : f(x) \leq c\} = \overline{\{x : f(x) < c\}} \quad (2.1)$$

$$\{x : c \leq f(x)\} = \overline{\{x : c < f(x)\}}. \quad (2.2)$$

Theorem 2.1. *Under assumption (A1), the estimator $\hat{\mathbf{F}}_n(c_1, c_2)$ converges strongly to $F(c_1, c_2)$ if the function f satisfies condition (2.1) at $c = c_2$ and condition (2.2) at $c = c_1$. Moreover, (2.1) and (2.2) are necessary in the following sense:*

1. *Suppose that x_0 is a point such that there exists a neighbourhood $B_\delta(x_0)$ and a subsequence n_k such that $\hat{f}_{n_k}(x) > f(x)$ for all $x \in B_\delta(x_0)$. If $\hat{\mathbf{F}}_n(c_1, c_2)$ is consistent, then (2.1) must hold at x_0 for $c = c_2$ in the sense that $x_0 \notin \{x : f(x) \leq c_2\} \setminus \overline{\{x : f(x) < c_2\}}$.*
2. *Suppose that x_0 is a point such that there exists a neighbourhood $B_\delta(x_0)$ and a subsequence n_k such that $\hat{f}_{n_k}(x) < f(x)$ for all $x \in$*

$B_{\mathcal{D}}(x_0)$. If $\hat{\mathbf{F}}_n(c_1, c_2)$ is consistent, then (2.2) must hold at x_0 for $c = c_1$ in the sense that $x_0 \notin \{x : c_1 \leq f(x)\} \setminus \overline{\{x : c_1 < f(x)\}}$.

Example 1: Let $\mathcal{D} = [-2, 2]^2 \subset \mathbb{R}^2$ with $f(x) = |x|$ and $F(-\infty, 1) = \{x : f(x) \leq 1\}$, the disc with radius one centred at the origin. Suppose U_1, \dots, U_n are independent and identically distributed random variables from the uniform distribution on $[-1, 1]^2$, and let \bar{U}_n denote their bivariate sample mean. Then $\hat{f}_n(x) = |x - \bar{U}_n|$ converges uniformly to $f(x)$ on \mathcal{D} , and $f(x)$ satisfies (2.1) and (2.2) at $p = 1$. Therefore, $\hat{\mathbf{F}}_n(-\infty, 1)$ and $\hat{\mathbf{F}}_n(1, 1)$ are consistent for $F(-\infty, 1)$ and $F(1, 1) = \{x : |x| = 1\}$. In this case, the Hausdorff distance $\rho(\hat{\mathbf{F}}_n(-\infty, 1), F(-\infty, 1)) = \rho(\hat{\mathbf{F}}_n(1, 1), F(1, 1)) = |\bar{U}_n|$ converges to zero almost surely, by the strong law of large numbers.

3. CONFIDENCE REGIONS

Now, assume that the estimating functions \hat{f}_n satisfy assumption (A2) for some compact window $\mathcal{W} \subset \mathcal{D}$. The confidence regions for the sets (1.1) restricted to \mathcal{W} can be obtained as follows. Let q_1 and q_2 be the quantiles of the process $\sup_{x \in \mathcal{W}} \mathbb{Z}(x)$ such that

$$\begin{aligned} \Pr\left(\sup_{x \in \mathcal{W}} \mathbb{Z}(x) \leq q_1\right) &= 1 - \alpha, \\ \Pr\left(\sup_{x \in \mathcal{W}} |\mathbb{Z}(x)| \leq q_2\right) &= 1 - \alpha. \end{aligned}$$

Then the sets

$$\begin{aligned} &\left\{x \in \mathcal{W} : \hat{f}_n(x) \leq c + q_1/\sqrt{n}\right\}, \\ &\left\{x \in \mathcal{W} : c_1 - \frac{q_2}{\sqrt{n}} \leq \hat{f}_n(x) \leq c_2 + \frac{q_2}{\sqrt{n}}\right\}, \end{aligned} \quad (3.1)$$

form $100(1 - \alpha)\%$ confidence regions for $\mathcal{W} \cap F(-\infty, c)$ and $\mathcal{W} \cap F(c_1, c_2)$, respectively (where $-\infty < c_1 \leq c_2 < \infty$ and $c \in \mathbb{R}$). Note that the random variables $\sup_{x \in \mathcal{W}} |\mathbb{Z}(x)|$ and $\sup_{x \in \mathcal{W}} \mathbb{Z}(x)$ are well-defined because \mathcal{W} is compact and \mathbb{Z} has continuous sample paths. This also implies that we may use $\max_{x \in \mathcal{W}} |\mathbb{Z}(x)|$ and $\max_{x \in \mathcal{W}} \mathbb{Z}(x)$ to calculate the quantiles, which is computationally

easier. In what follows, we assume that $\mathcal{W} = \mathcal{D}$, where \mathcal{D} is compact, unless otherwise stated.

Recall that the coverage of a confidence region refers to the probability with which it covers the quantity of interest. For a $100(1 - \alpha)\%$ confidence region, it is typically considered ideal if the coverage is as close to $100(1 - \alpha)\%$ without going over. Such a confidence region is preferred because it is *conservative*, in the sense that it will never under quantify the variability of the estimator. Let us first show that our approach yields such a confidence region.

$$\begin{aligned} &\left\{F(-\infty, c) \subset \{\hat{f}_n(x) \leq c + q_1/\sqrt{n}\}\right\}^c \\ &= \left\{\hat{f}_n(x) > c + q_1/\sqrt{n} \exists x \in F(-\infty, c)\right\}. \end{aligned}$$

Now, let $\mathbb{Z}_n(x) = \sqrt{n}(\hat{f}_n(x) - f(x))$. Then

$$\begin{aligned} &\Pr\left(\hat{f}_n(x) > c + q_1/\sqrt{n} \exists x \in F(-\infty, c)\right) \\ &= \Pr\left(\mathbb{Z}_n(x) > \sqrt{n}(c - f(x)) + q_1 \exists x \in F(-\infty, c)\right) \\ &\leq \Pr(\mathbb{Z}_n(x) > q_1 \exists x \in F(-\infty, c)) \\ &\leq \Pr(\mathbb{Z}_n(x) > q_1 \exists x), \end{aligned} \quad (3.2)$$

and taking the limit in n , the latter quantity is less than or equal to α by definition of q_1 . It therefore follows that

$$\Pr(F(-\infty, c) \subset \{\hat{f}_n(x) \leq c + q_1/\sqrt{n}\}) \geq 1 - \alpha,$$

as required, asymptotically. A similar approach works for the case $F(c_1, c_2)$.

The above calculation is illuminating for several reasons:

- Notice that the consistency conditions play no role in the design of the confidence region. Indeed, the confidence region functions as intended even if consistency is violated (see e.g. Example 3).
- The smoothness and variability of the field \mathbb{Z} determines the “size” of the confidence set, which may not be uniform over \mathcal{W} . In fact, the larger the window \mathcal{W} is chosen, the wider

the confidence set is. Furthermore, going from the penultimate to the ultimate line in (3.2) we note that the bounding term $\sup \mathbb{Z}(x)$ could be replaced with many other upper bounds, including $\sup_{x \in F(-\infty, c)} \mathbb{Z}(x)$, and asymptotically with $\sup_{x \in F(c, c)} \mathbb{Z}(x)$.

- The calculations in (3.2) give also some ideas on the power of the methodology. In this context, statistical power would indicate an ability to also recognize the regions which should not be in $F(c_1, c_2)$ with high probability. For $x \in F(-\infty, c)^c$, the quantity $\sqrt{n}(c - f(x))$ converges to $-\infty$, and therefore these locations should be easily picked up by the confidence region for large enough sample sizes.
- The confidence regions are conservative, in that the upper bounds in (3.2) mean that over-coverage is possible, even asymptotically. We explore the extent of this via simulations in Section 4.

Finally, we note that our methods are straightforward to implement, and, if exact limiting distributions cannot be found, the confidence sets may be estimated using bootstrap methods, either parametric or nonparametric [5, for additional details on bootstrap methodologies]. In what follows, we used the parametric bootstrap in Example 5 and the nonparametric bootstrap in Example 6.

Example 2: To build 95% confidence regions for Example 1: $F(1, 1) = \{x : |x| = 1\}$ with $\mathcal{W} = \mathcal{D} = [-2, 2]^2$, we calculate $\mathbb{Z}_n(x) = \sqrt{n}(|x - \bar{U}_n| - |x|)$, where \bar{U}_n is the average of n independent Uniform $[-1, 1]^2$ random variables. Clearly, $\mathbb{Z}_n(x)$ has continuous sample paths. Also, $|\mathbb{Z}_n(x)| \leq \sqrt{n}|\bar{U}_n|$ for all x and, since this is realized at $x = 0$, we obtain that $\sup_{x \in \mathcal{D}} |\mathbb{Z}_n(x)| = \sqrt{n}|\bar{U}_n|$. The limiting distribution is therefore $\sqrt{3^{-1}(Z_1^2 + Z_2^2)}$ where Z_1, Z_2 are independent standard normal variables, and it therefore follows that $q_2 = 1.41$.

4. SIMULATION STUDY

In (3.2), we have shown that the confidence regions (3.1) cover the true level set at least

Table 1: Empirical coverage probabilities for 95% confidence region. The standard error due to Monte Carlo sampling is 0.0022.

	$n = 25$	$n = 100$	$n = 1000$
(A)	95.26	94.83	95.12
(B)	94.79	95.34	95.47
(C)	97.81	98.16	98.45

$100(1 - \alpha)\%$ of the time, for sufficiently large n . Here, we use Monte Carlo simulations to calculate the empirical coverage probabilities for various examples and sample sizes. The goal is to understand the actual behaviour of the methods, as well as the amount of over-coverage one could expect to see in practice. The cases we consider are (A) Example 2 of Section 3, and (B) Example 3 and (C) Example 4 described below. For each of the three examples, we calculate the $100(1 - \alpha)\%$ confidence region with $\mathcal{W} = \mathcal{D}$ and estimate the coverage. We consider the following sets:

		$\mathcal{W} = \mathcal{D}$
(A)	$F(1, 1)$	$[-2, 2]^2$
(B)	$F(-\infty, 1)$	$[-1, 2]$
(C)	$F(0, 1)$	$[-2, 2]^2$

The simulations were done in MATLAB, on a discretized domain \mathcal{D} . Because the discretization introduces some error into the calculations [22], we selected a large lattice and calibrated it to give accurate results as follows. Suppose that $\mathcal{D} \subset \mathbb{R}^2$ and $f(x) = |x| - 0.5$ with $\hat{f}_n(x) = |x| - \bar{U}_n$, where \bar{U}_n is the average of n independent samples from a uniform distribution on $[0, 1]$. The confidence regions for $F(0, 0)$ are exact because of the special separable form of the function \hat{f}_n (modulo the sample size approximations). For $n = 1000$ and lattices with $m = 200, 400, 600, 800, 1000$, the empirical coverage probabilities for the 95% confidence region were 95.10, 95.02, 95.30, 95.30, and 94.96s, respectively; the standard error due to Monte Carlo sampling was 0.003. From here,

Table 2: Empirical coverage probabilities for 90% confidence region; the standard error due to Monte Carlo sampling is 0.0030.

	$n = 25$	$n = 100$	$n = 1000$
(A)	89.87	89.70	90.17
(B)	88.49	90.07	90.17
(C)	96.10	96.09	96.71

we selected $m = 600$ for simulations.

Example 3: Let $\mathcal{D} = [-1, 2] \subset \mathbb{R}$ and $h(x) = |x - 0.5| - 0.5$. Let $\hat{f}_n(x) = \hat{p}_n|h(x)| + (1 - \hat{p}_n)h(x)$, where $n\hat{p}_n$ is a binomial random variable with parameters n and $p = 1/2$. Set the true function $f(x) = E[\hat{f}_n(x)] = \max(h(x), 0)$ and $F(-\infty, 0) = \{x : f(x) \leq 0\} = [0, 1]$. Notice that f does not satisfy condition (2.1) at $c = 0$. Indeed, if $\hat{p}_n > 0.5$ then $\hat{F}_n(-\infty, 0) = \{0, 1\}$ and otherwise $\hat{F}_n(-\infty, 0) = [0, 1]$. Clearly, consistency does not hold, however, the confidence regions will still behave as expected.

Example 4: Suppose that $\mathcal{D} = [-2, 2]^2$ and $f(x) = \beta^T \tilde{x}$, where $\beta^T = (0.5, 1, 2, -3, 1)$ and $\tilde{x} = (1, x_1, x_2, x_1x_2, x_1^2)^T$. This is estimated by the regression function $\hat{f}_n(x) = \hat{\beta}_n^T \tilde{x}$, where $\hat{\beta}_n$ is normally distributed with mean β and variance Σ/n with $\Sigma_{ii} = 1$ and $\Sigma_{ij} = 0.2$ for $i \neq j$. The quantiles of the fluctuation field $\sup_{x \in \mathcal{D}} |\mathbb{Z}(x)|$ were found empirically. The set being estimated is $F(0, 1)$. We show examples of the confidence regions for $n = 25, 100$ and 1000 in Figure 2. We refer to Example 5 for a more detailed regression example.

Tables 1 and 2 show the results for $B = 10,000$ Monte Carlo simulations. Although the confidence regions are conservative, (A) and (B) both show almost exact coverage. In (C), the variability of the fluctuation field is reflected in the size of the confidence sets. The effect is compounded because the function f is relatively flat in the neighbourhood of $F(0, 1)$.

5. EXAMPLES AND APPLICATIONS

5.1 Covariate Domain Estimation

Least squares regression is a well-known statistical tool ubiquitous across the sciences. In regression analysis, the expected response is modeled as a function of covariates, i.e. we model $E[Y|X = x]$ where Y is the response variable and x is the vector of covariates. Given $f(x) = E[Y|X = x]$, the various sets (1.1) describe the domain of covariates for which the mean response lies within a specified target range. Although we focus on the linear regression model, the method can be easily extended to most generalized linear models. We thus define $f(x) = \beta^T \tilde{x}$, with $\tilde{x} = \tilde{x}(x)$ denoting some continuous function of the covariates. We assume that the covariates are continuous and lie in $\mathcal{D} \subset \mathbb{R}^d$, with $p + 1 \geq d$ denoting the number of regression variables. For example, if $f(x) = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_1x_2$, then $d = 2, p = 3$ and $\tilde{x}^T = \tilde{x}^T(x_1, x_2) = (1, x_1, x_2, x_1x_2)$.

Consider the estimating function

$$\hat{f}_n(x) = \hat{\beta}_n^T \tilde{x},$$

where n is the number of observations. Since \tilde{x} is continuous, the image of \mathcal{D} under \tilde{x} is compact. It follows that $\hat{f}_n(x)$ converges uniformly to $f(x)$ as long as $\hat{\beta}_n$ are consistent estimators, and therefore, \hat{f}_n satisfies assumption (A1). The conditions (2.1) and (2.2) need to be checked on a case by case basis. For example, for $f(x) = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_1x_2$, both conditions are satisfied for any value of c , as long as at least one of $\beta_1, \beta_2, \beta_3$ is non-negative.

Next, let $\mathbb{Z}_n(x) = \sqrt{n}(\hat{f}_n(x) - f(x))$. If $\hat{\beta}_n$ is asymptotically normal, we have that $\sqrt{n}(\hat{\beta}_n - \beta) \Rightarrow Z$, where Z is multivariate normal with mean zero and variance Σ . If unknown, Σ is estimated using standard regression methods. Since \tilde{x} is continuous, it follows that $\mathbb{Z}_n(x)$ converges weakly in $C(\mathcal{D})$ to a continuous, mean zero Gaussian field, $\mathbb{Z}(x) = Z^T \tilde{x}$, with covariance structure given by $\text{cov}(\mathbb{Z}(x), \mathbb{Z}(x')) = \tilde{x}^T \Sigma \tilde{x}$. Therefore, \hat{f}_n satisfies assumption (A2), and confidence sets

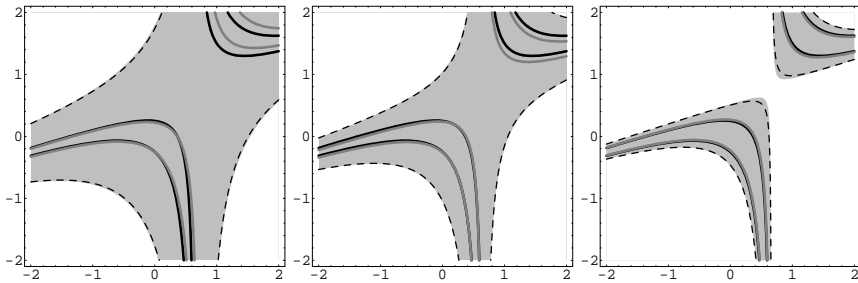


Figure 2: For each of $n = 25, 100, 1000$ (from left to right), the boundary of the true set $F(0, 1)$ (black) and the boundary estimate $\hat{F}_n(0, 1)$ (dark grey) are shown along with the 90% confidence region (light grey). The boundary of the modified 90% confidence region is also shown (dashed).

may be formed as described above.

In [3], this problem was considered for a logistic regression in the context of effective dose estimation, and this approach was later studied numerically in [12]. The idea we present here and the approach of [3] are similar in that they both obtain upper bounds on the supremum of the fluctuation process. The upper bounds of [3] are derived via a similar idea to that used for Scheffé's bounds for simultaneous confidence intervals [21]. We refer to [3] and [12] for more details. The general methodology described in Section 3 does not specify the best way in which to obtain these quantiles. The approach of [3] is one which does work for general settings of parametric regression problems, but these bounds are very conservative, and therefore result in much over coverage for the methodology [see 12]. In the example below, we show a more direct approach of estimating the quantiles of $\sup_{x \in \mathcal{D}} \mathbb{Z}(x)$.

Example 5: We illustrate the method on the data set `trees` available with R [20]. Here, girth (in inches), height (in feet) and volume (in cubic feet) of timber were recorded for 31 felled black cherry trees. Set $x_1 =$ girth and $x_2 =$ height. Fitting the model $E[\log Y|x] = \beta_0 + \beta_1 \log x_1 + \beta_2 \log x_2$, we obtain estimates $\hat{\beta}_0 = -6.63$ ($p\text{-value} = 5.1e -$

09), $\hat{\beta}_1 = 1.98$ ($p\text{-value} < 2e - 16$) and $\hat{\beta}_2 = 1.12$ ($p\text{-value} = 7.8e - 06$). The estimates are not far from the formula $\text{volume} = \text{height} \times \text{girth}^2/4\pi$.

Set $\mathcal{D} = [5, 25] \times [50, 100]$, and suppose that we are interested in the domain of covariates for which the log-volume is at least $\log 30$ (≈ 3.4), that is,

$$\begin{aligned} F(-\log 30, \infty) &= \{x : E[\log Y|x] \geq \log 30\} \\ &= \{x : -\beta_0 - \beta_1 \log x_1 - \beta_2 \log x_2 \leq -\log 30\}. \end{aligned}$$

Figure 3 shows the estimator $\hat{F}_n(-\log 30, \infty) = \{x : \hat{f}_n(x) \leq -\log 30\}$, where $\hat{f}_n(x) = \hat{\beta}^T \tilde{x}$ and $\tilde{x}^T = (-1, -\log x_1, -\log x_2)$. Note that \tilde{x} is continuous on \mathcal{D} .

The true function $f(x) = -\beta_0 - \beta_1 \log x_1 - \beta_2 \log x_2$ is strictly decreasing in both x_1 and x_2 , and therefore, it satisfies condition (2.1) at $c = -\log 30$, or for any other choice of c . Condition (2.2) is also satisfied, although we do not require it here. It follows that the set $\hat{F}_n(-\log 30, \infty)$ is consistent for $F(-\log 30, \infty)$.

The 95% confidence set for $F(-\log 30, \infty)$ is $\{x : \hat{f}_n(x) \leq -\log 30 + q_1/\sqrt{31}\}$, where q_1 is the value such that $\text{pr}(\sup_{x \in \mathcal{D}} \mathbb{Z}(x) \leq q_1) = 0.95$. In this case, the fluctuation process is $\mathbb{Z}(x) = Z^T \tilde{x}$,

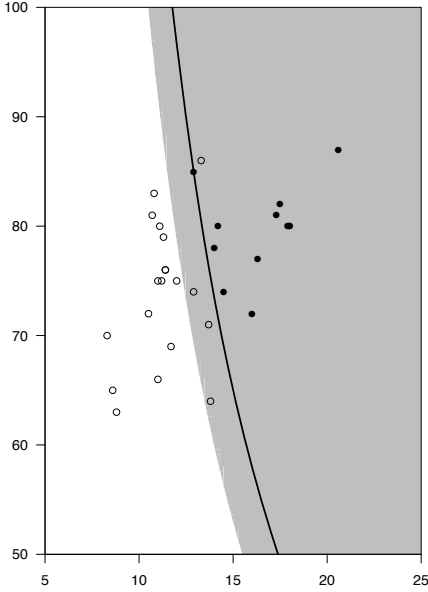


Figure 3: The estimated set $\hat{F}_n(-\log 30, \infty)$ (the boundary of this set is shown in black) with the 95 % confidence region (light gray). The data points are also shown as circles, with the filled in circles showing those data with volume ≥ 30 .

where $Z^T = (Z_0, Z_1, Z_2)$ is a mean-zero multivariate Gaussian. Under the normal linear model, a consistent estimator of the covariance matrix of Z is $\hat{\Sigma}$ where

$$\begin{aligned}\hat{\Sigma} &= n\hat{\sigma}^2(X'X)^{-1} \\ &= 31 \times \begin{bmatrix} 0.6397 & 0.0208 & -0.1601 \\ 0.0208 & 0.0056 & -0.0081 \\ -0.1601 & -0.0081 & 0.0418 \end{bmatrix},\end{aligned}$$

where X is the design matrix of the regression. The supremum of \mathbb{Z} must occur on one of the corners of \mathcal{D} ,

$$\sup_{z \in \mathcal{D}} \mathbb{Z}(x) = \max_{i=1,2, j=1,2} \{\mathbb{Z}(a_i, b_j)\},$$

for $a_1 = \log 5, a_2 = \log 25, b_1 = \log 50$ and $b_2 = \log 100$. We estimate the quantile via Monte

Carlo sampling from a multivariate Gaussian with variance $\hat{\Sigma}$, to obtain

$$\left\{x : -\beta_0 - \beta_1 \log x_1 - \beta_2 \log x_2 \leq -\log 30 + \frac{1.28}{\sqrt{31}}\right\}$$

as the approximate 95% confidence region. The resulting set is shown in Figure 3.

5.2 Estimating the Boundary of a Density Region

Here, our goal is to identify the set $F(c, c) = \{x : f(x) \geq c\}$ for the unknown density f . This problem is closely related to clustering, in the sense that for certain cutoff values c , the set $F(c, c)$ can be used to identify both the number of clusters and their centres.

Example 6: Consider the problem of estimating the level set of density function. Suppose that f is a mixture density given by $f = 0.5g_1 + 0.5g_2$, where g_1 and g_2 are both bivariate Gaussian densities. Figure 4 shows two examples of such density functions with different degrees of separation between the mixture components g_i . Using a random sample of size $n = 1000$, we estimate the level set $F(0.055, 0.055) = \{x : f(x) = 0.055\}$ for both examples. Wider confidence regions in Figure 5 (left) indicate that the lack of separation between the mixture components leads to higher variability (equivalently, less accuracy) of the estimator.

The drastic difference in Figure 5 (top vs. bottom) implies that confidence regions can be effectively used to characterize and visualize the variability and accuracy of level set estimators.

This idea closely relates to estimation of high intensity regions, where one is interested in finding “hot spots”, or regions where the estimated intensity crosses some threshold. Important examples include high/low vegetation growth or regions exhibiting high cancer rates [17] or probabilistic forecasting of extreme weather events in meteorology [14].

6. DISCUSSION

We present a method for graphically presenting the risk associated with replacing the unknown

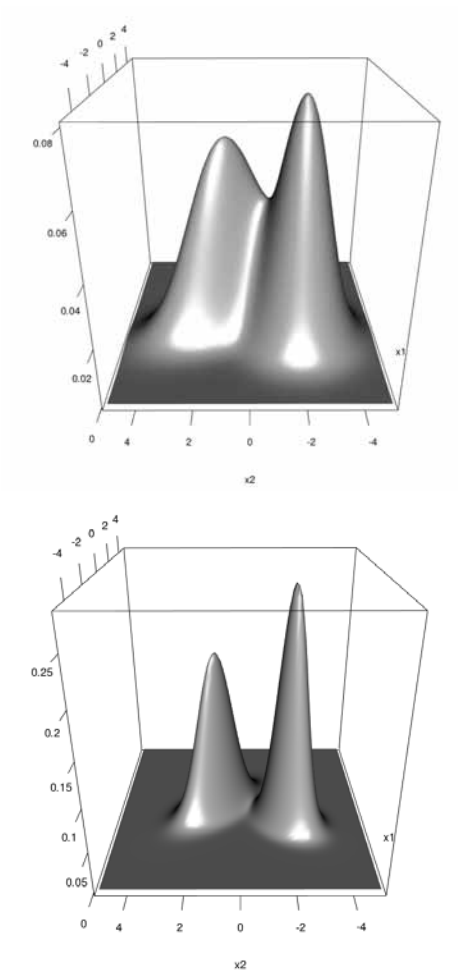


Figure 4: Two Gaussian mixture densities : f (top) and \tilde{f} (bottom), as defined in the text.

function with its estimator when computing a level set. The method is appropriately conservative, and we have studied the amount of overcoverage in several examples through simulations. We also give some specific examples based on simulated and real data.

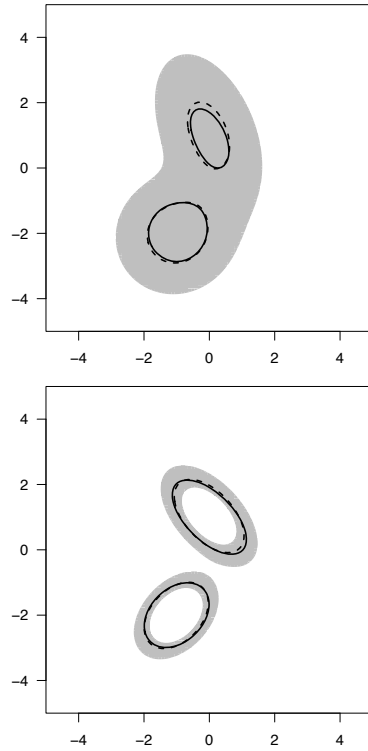


Figure 5: The boundary of true level set $F(0.055, 0.055)$ (solid line) and the boundary of the estimated level set $\hat{F}_n(0.055, 0.055)$ (dashed line) for densities f (top) and \tilde{f} (bottom) in Figure 4. The level set estimates are based on a random sample of size $n = 1000$. Visually, both the level sets and their estimates appear quite similar. However, the 95% confidence regions (gray) reveal great variability in the accuracy of the two estimates.

APPENDIX

Technical details

Lemma 6.1. Suppose that f is continuous. Then

$$\begin{aligned} \{x : c_1 \leq f(x)\}^\varepsilon \cap \{x : f(x) \leq c_2\}^\varepsilon \\ = \{x : c_1 \leq f(x) \leq c_2\}^\varepsilon. \end{aligned}$$

Proof. Suppose y is in the set $\{x : c_1 \leq f(x)\}^\varepsilon \cap \{x : f(x) \leq c_2\}^\varepsilon$ but not in the set $\{x : c_1 \leq f(x) \leq c_2\}$. Then one of two possibilities exists: Either $f(y) < c_1$ or $f(y) > c_2$. The argument for both cases is the same, so we present only the first instance.

Assume then that y is such that $f(y) < c_1$. By definition of y , there exists an x_1 such that $c_1 \leq f(x_1)$ and $y \in B_\varepsilon(x_1)$, or an x_2 such that $f(x_2) \leq c_2$ and $y \in B_\varepsilon(x_2)$. In the first setting, since f is continuous, there also exists a z such that $c_1 \leq f(z) \leq c_2$ and $d(y, z) \leq \varepsilon$. For example, one such z must fall on the line between y and x_1 , which would clearly satisfy $|y - z| \leq \varepsilon$. A similar argument shows that if $y \in B_\varepsilon(x_2)$, then there exists a z such that $c_1 \leq f(z) \leq c_2$ and $|y - z| \leq \varepsilon$. It follows that $y \in B(z, \varepsilon)$. This proves that

$$\begin{aligned} \{x : c_1 \leq f(x)\}^\varepsilon \cap \{x : f(x) \leq c_2\}^\varepsilon \\ \subset \{x : c_1 \leq f(x) \leq c_2\}^\varepsilon. \end{aligned}$$

Containment in the other direction is immediate, completing the proof. \square

Proof of Theorem 2.1. The proof here is similar to that of [16]. Without loss of generality, we may assume that \mathcal{D} is compact. If $f : \mathcal{D} \mapsto \mathbb{R}$ is a continuous function satisfying the conditions of the theorem, then

$$\begin{aligned} \tilde{\varphi}(\pm\varepsilon) &= \rho(\{x : f(x) \leq c_2\}, \{x : f(x) \leq c_2 \pm \varepsilon\}) \\ \varphi(\pm\varepsilon) &= \rho(\{x : c_1 \leq f(x)\}, \{x : c_1 \pm \varepsilon \leq f(x)\}) \end{aligned}$$

are all continuous for ε near zero, and moreover, they both converge to zero as $\varepsilon \rightarrow 0$. Now, by (A1), we know that \hat{f}_n converges uniformly to f with probability one. Let

$$\eta_n = \sup_{x \in \mathcal{D}} |f(x) - \hat{f}_n(x)|,$$

and also define

$$\varepsilon_n = \max\{\varphi(\eta_n), \varphi(-\eta_n), \tilde{\varphi}(\eta_n), \tilde{\varphi}(-\eta_n)\}$$

which converges to zero as $n \rightarrow \infty$ almost surely. We will next show that $\rho(\{x : c_1 \leq \hat{f}_n(x) \leq$

$c_2\}, \{x : c_1 \leq f(x) \leq c_2\}) \leq \varepsilon_n$. To this end

$$\begin{aligned} \{x : c_1 \leq f(x) \leq c_2\} \\ \subset \{x : f(x) \leq c_2 - \eta_n\}^{\tilde{\varphi}(-\eta_n)} \\ \subset \{x : \hat{f}_n(x) \leq c_2\}^{\tilde{\varphi}(-\eta_n)} \subset \{x : \hat{f}_n(x) \leq c_2\}^{\varepsilon_n}. \end{aligned}$$

Repeating in the other direction, we obtain

$$\begin{aligned} \{x : c_1 \leq f(x) \leq c_2\} \\ \subset \{x : c_1 + \eta_n \leq f(x)\}^{\varphi(\eta_n)} \\ \subset \{x : c_1 \leq \hat{f}_n(x)\}^{\varphi(\eta_n)} \subset \{x : c_1 \leq \hat{f}_n(x)\}^{\varepsilon_n}. \end{aligned}$$

and hence, by Lemma 6.1,

$$\begin{aligned} \{x : c_1 \leq f(x) \leq c_2\} \\ \subset \{x : c_1 \leq \hat{f}_n(x) \leq c_2\}^{\varepsilon_n}. \quad (\text{A-1}) \end{aligned}$$

For the other direction,

$$\begin{aligned} \{x : c_1 \leq \hat{f}_n(x) \leq c_2\} &\subset \{x : f(x) \leq c_2 + \eta_n\} \\ &\subset \{x : f(x) \leq c_2\}^{\varepsilon_n}. \end{aligned}$$

A similar argument shows that $\{x : c_1 \leq \hat{f}_n \leq c_2\} \subset \{x : c_1 \leq f(x)\}^{\varepsilon_n}$, from which it follows

$$\{x : c_1 \leq \hat{f}_n(x) \leq c_2\} \subset \{x : c_1 \leq f(x) \leq c_2\}^{\varepsilon_n}$$

by Lemma 6.1. Together with (A-1) this proves the result.

To address necessity, suppose that there exists a neighbourhood of x_0 , $B_\delta(x_0)$, and a subsequence n_k such that $\hat{f}_{n_k}(x) < f(x)$ for all $x \in B_\delta(x_0)$. Assume also that $x_0 \in \{x : c_1 \leq f(x)\} \setminus \{x : c_1 < f(x)\}$. In particular, this implies that (2.2) is not satisfied, and hence there exists an $\varepsilon > 0$ such that $\rho(x_0, \{x : c_1 < f(x) \leq c_2\}) > \varepsilon$. It follows that $\rho(\hat{F}_{n_k}(c_1, c_2), F(c_1, c_2)) > \min(\varepsilon, \delta) > 0$, proving the result. A similar argument proves the other claim. \square

ACKNOWLEDGMENTS

HJ would like to thank NSERC for funding support.

REFERENCES

- [1] B. N. Bekele and P. F. Thall. Dose-finding based on multiple toxicities in a soft tissue sarcoma trial. *J. Amer. Statist. Assoc.*, 99(465): 26–35, 2004.
- [2] P. Billingsley. *Convergence of probability measures*. John Wiley & Sons Inc., New York, 1968.
- [3] W. Carter, V. Chinchilli, J. Wilson, E. Campbell, F. Kessler, and R. Carchman. An asymptotic confidence region for the ED100p from the logistic response surface for a combination of agents. *The American Statistician*, 40(2): 124–128, 1986.
- [4] A. Cuevas, W. González-Manteiga, and A. Rodríguez-Casal. Plug-in estimation of general level sets. *Aust. N. Z. J. Stat.*, 48(1): 7–19, 2006.
- [5] A. Davison and H. D. *Bootstrap methods and their application*. Cambridge University Press, 1997.
- [6] H. Jankowski, X. Ji, and L. Stanberry. A random set approach to confidence regions with applications to the effective dose with combinations of agents. *Statistics in Medicine*, 2014. To appear.
- [7] H. Jankowski, V. Sabelnykova, and J. Sheriff. Estimating the wintering location of the wood thrush. Technical report, York University, 2011.
- [8] H. Jankowski and L. Stanberry. Confidence regions for means of random sets using oriented distance functions. *Scandinavian Journal of Statistics*, 39(2): 340–357, 2012.
- [9] H. K. Jankowski and L. I. Stanberry. Expectations of random sets and their boundaries using oriented distance functions. *Journal of Mathematical Imaging and Vision*, 36(3): 291–303, 2010.
- [10] A. P. Korostel'ev and A. B. Tsybakov. *Minimax theory of image reconstruction*, volume 82 of *Lecture Notes in Statistics*. Springer-Verlag, New York, 1993.
- [11] H. Kunita. *Stochastic flows and stochastic differential equations*, volume 24 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1990.
- [12] J. Li, E. Nordheim, C. Zhang, and C. Lehner. Estimation and confidence regions for multi-dimensional effective dose. *Biometrical Journal*, 50(1): 110–122, 2008.
- [13] E. Mammen and W. Polonik. Confidence sets for level sets. *Journal of Multivariate Analysis*, 122: 202–214, 2013.
- [14] C. Mass, S. Joslyn, J. Pyle, P. Tewson, T. Gneiting, A. Raftery, J. Baars, J. M. Sloughter, D. Jones, and C. Fraley. PROBCAST: A web-based portal to mesoscale probabilistic forecasts. *Bulletin of the American Meteorological Society*, 90: 1009–1014, 2009.
- [15] G. Matheron. *Random sets and integral geometry*. John Wiley & Sons, New York-London-Sydney, 1975. With a foreword by Geoffrey S. Watson, Wiley Series in Probability and Mathematical Statistics.
- [16] I. S. Molchanov. A limit theorem for solutions of inequalities. *Scandinavian Journal of Statistics*, 25: 235–242, 1998.
- [17] P. Nguyen, P. Brown, and S. J.E. Mapping cancer risk in Southwestern Ontario with changing census boundaries. *Biometrics*, 68, 2012.
- [18] D. Nolan. The excess-mass ellipsoid. *J. Multivariate Anal.*, 39(2): 348–371, 1991.
- [19] W. Polonik. Measuring mass concentrations and estimating density contour clusters—an

excess mass approach. *Ann. Statist.*, 23(3): 855–881, 1995.

- [20] R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2008. URL <http://www.R-project.org>.
- [21] H. Scheffé. A method for judging all contrasts in the analysis of variance. *Biometrika*, 40: 87–104, 1953.
- [22] J. Serra. *Image analysis and mathematical morphology*. Academic Press Inc., London, 1984.
- [23] P. F. Thall and J. D. Cook. Dose-finding based on efficacy-toxicity trade-offs. *Biometrics*, 60(3): 684–693, 2004.
- [24] P. F. Thall, R. E. Millikan, P. Mueller, and S.-J. Lee. Dose-finding with two agents in Phase I oncology trials. *Biometrics*, 59(3): 487–496, 2003.
- [25] S. R. S. Varadhan. *Stochastic processes*, volume 16 of *Courant Lecture Notes in Mathematics*. Courant Institute of Mathematical Sciences, New York, 2007.

ABOUT THE AUTHORS

1. Hanna Jankowski is Associate Professor in the Department of Mathematics and Statistics at York University in Toronto, Canada.
2. Larissa Stanberry is Biostatistician at Predictive Analytics, Seattle Children’s Hospital, Seattle, USA.

VISUALIZATION OF TRAVEL TIME BY TRANSPORTATION SERVICE WITH CONCENTRIC CIRCLES

Michio SHIRAISHI¹ and Aya INOUE² and Mikio SHINYA¹

¹Toho University, Japan ²U-DOM Co., Ltd., Japan

ABSTRACT: Time needed to travel from a location to another is not usually in proportion to the geodesic distance between two locations, because many transportation services such as subways or trains are available in large cities in the world. A variety of web sites provides train route finders which give paths and durations to travel between two locations, but when we want to know the places where we can reach within given period, such route finders are not sufficient.

We developed a web application which visualizes travel time from a given location to any other locations using transportation services by overlaying colored regions on a map. A user inputs the location where the travel starts, and then the system queries the online Internet service about the travel time from the nearest station to other stations. Assuming that users walk straight from stations, we can represent these regions which indicate time ranges as concentric circles centered at the station. We employed WebGL to efficiently draw regions in web browsers using a stencil buffer. We confirm that users can easily grasp the distribution of the travel time.

Keywords: Visualization, travel time, transportation service, WebGL

1. INTRODCUTION

Public transportation service is essential in many metropolises. Especially, railroads and subways are free from traffic jams, and various train routes have formed a complex network in many large cities. When we want to travel from one location to another using such services, it is convenient to search routes with online services, often provided by portal sites. They are useful when the destination is already known, but the travel time needed between two locations cannot be easily found by such services.

Visualizing spatial distribution of travel time from a location is useful for residents who use the services in their daily life as well as tourists who are not familiar with such services. For example, when we search hotels for a trip, we often compare their accessibility.

Map visualization is commonly used to explore the space distribution of data[1]. We often draw glyphs on a map to visualize spatial

patterns to see the geographical distribution[2]. One of the important applications of map visualization is the analysis of spatial distribution of travel time, which is important in the field of transport geography. For example, Mesbar *et al.* discussed the spatial visualization of travel time to the central business district in Melbourne using the tram network[3]. Although they succeed in visualizing the result of operation data analysis, it is restricted that the travel time to one location in interest.

In this paper, we propose a method that visualizes distributions of travel time from a given location of departure interactively. In current implementation, the web application queries the transportation information service and renders the results on Google Maps. Actually some web sites [4][5][6] have the same functionalities as our application, but our implementation can render more comprehensive geometries with the help of WebGL.

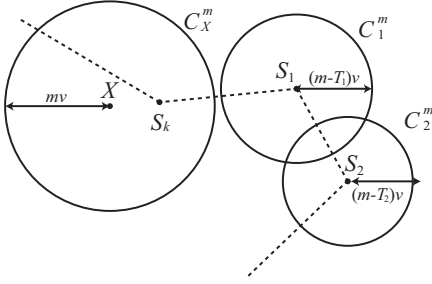


Figure 1: The regions reachable within m minutes. The dashed line indicates the route of public transportation service.

2. PROPOSED METHOD

2.1 Basic Algorithm

In this section, we introduce a method to region reachable within m minutes A^m from the given location. The input is the address or the name of landmark from where travel starts. Then, the region reachable in 10 minutes, one reachable in 10 to 20 minutes, and so forth, are displayed on the map. Each region is rendered with its own translucent color.

Let X be the start location represented by a pair of latitude and longitude. Assume that the application has the list of train station locations $S_i (i = 1, 2, \dots, n)$.

Firstly, the nearest station from X , $S_k (k \in \{1, 2, \dots, n\})$, is searched using pre-built kd-tree of station locations $\{S_i\}$. Once S_k is found, the time t_k needed to walk from X to S_k is calculated using a standard walking speed v . When a traveler walks from X , the region reachable in m minutes becomes the circle C_X^m centered at X as shown in Figure 1. The diameter of this circle corresponds to the geodesic distance the user can walk straight from X in m minutes, mv .

Next, the travel time T_i from X to $S_i (i \neq k)$ is calculated as the sum of t_k and the travel time by the train from the station S_k to S_i . For each station S_i such that $T_i \leq m$, the circle C_i^m centered at X has the diameter of $(m - T_i)v$.

A^m is the union of C_X^m and all C_i^m as shown in

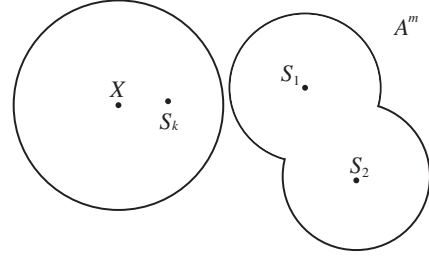


Figure 2: The region reachable within m minutes.

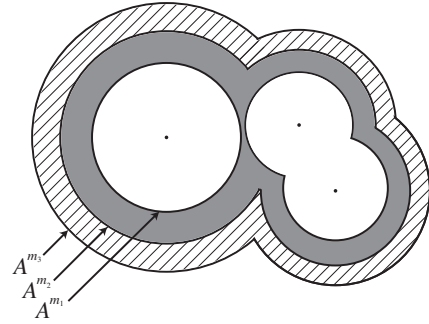


Figure 3: The regions reachable within m_1 minutes, m_2 minutes and m_3 minutes.

the Figure 2. We can define A^{m_1}, A^{m_2}, \dots for specified values of m ($m_1 < m_2 < \dots$) as shown in the Figure 3. Obviously, $A^{m_{j+1}}$ includes A^{m_j} , and the region reachable in m_j to m_{j+1} minutes is $A^{m_{j+1}} - A^{m_j}$.

2.2 Suggestion of meeting point

In section 2.1, we discussed traveling time from a start location, but we can apply the method for the case of multiple start locations. For instance, consider the situation when multiple persons should meet somewhere in the middle of their current locations. Obviously, it is preferable that the sum of travel time to a meeting point for each person is smaller. In addition to that, it is better if travel time for each person does not differ very much.

Let the travel time of a person j to a station s_i be $t_i(j)$. Then, for example, we can define c_i as the convenience of the station s_i by the following equation:

$$c_i = \sum_j t_i(j) - \min_i \sum_j t_i(j) + \text{Var}[t_i(j)],$$

where $\text{Var}[t_i(j)]$ is the variance of $t_i(j)$. Smaller c_i means that the station s_i is considered to be more convenient as a meeting point.

3. IMPLEMENTATION

The method described in the previous section is implemented as a web application which utilizes Google Maps[7]. The entry page has a text box in which a user inputs an address or the name of a landmark. Then the server creates a set of the station locations and their traveling time, and the web browser renders the regions based on the data returned from the server by using WebGL, a JavaScript API for utilizing low-level 3D graphics on web browsers[9]. In order to incorporate WebGL-based rendering in Google Maps, we can employ Google Maps Utility Library v3[8].

Note that all of web browsers do not support WebGL. In that case, polygons corresponding $A^{m_{j+1}} - A^{m_j}$ should be calculated on the server side because of the limitation of Google Maps JavaScript API v3. It is achieved by geometric set operations like JTS (Java Topology Suite) library[10], but it takes longer time than WebGL-based rendering.

The server processes an input in the following order and returns a set of the station locations.

1. The application converts an address or the name of a landmark to a pair of latitude and longitude by calling Google Geocoding API[11].
2. The application searches the nearest station from the local database.
3. The application accesses an online API service and puts the nearest station as a key to search the stations reachable within the

```

begin
  Enable stencil buffer
  foreach  $m_j$  (smaller to larger) do
    foreach  $S_i$  do
      if  $T_i - m_j > 0$  then
        Draw circle  $C_i^{m_j}$  (diameter =
           $(T_i - m_j)v$ ) for the pixels where
          the value of stencil buffer is 0.
        Set the value of stencil buffer to
          1 if drawn.
      end
    end
  end
end

```

Figure 4: Algorithm to render regions.

threshold minutes. The API returns a list of stations and their traveling time. We adopted Ekispart Web Service provided by Val Laboratory [12] to complete this task.

4. Finally, the server returns a set of station location coordinates and their traveling time from the start point to the web client.

The pseudocode of the rendering process in the client side is shown in Figure 4. Since the region reachable in m_j to m_{j+1} minutes is $A^{m_{j+1}} - A^{m_j}$, we render each A^{m_i} in the ascending order of m_i . This is achieved by using a stencil buffer to avoid render regions which are already rendered.

4. RESULTS

Fig. 5 shows the visualization of the travel time needed from Tokyo Station. The regions are separated in every 10 minutes. As shown from this figure, users can easily grasp the distribution of the travel time around the city. Although Shibuya (indicated as B) is closer to Tokyo Station (A) than Kawasaki (C), users can easily find that the travel times from Tokyo station are similar.

Fig. 6 shows candidates of a meeting point from Tokyo Station (indicated as A) and Shibuya Station(B). Since each station is colored according to the convenience of the station described

in 2.2, we can find Yotsuya (C) and Ochanomizu (D) may be a good meeting point.

5. CONCLUSIONS

We proposed a method to visualize travel time from a given location to other locations using transportation services by overlaying colored regions on a map. We implemented a web application which draws the regions on Google Maps using WebGL, which enables the efficient rendering of unions of circles using a stencil buffer. We confirm that users can easily grasp the distribution of the travel time around the city.

The travel time generally depends on departure time. We adopted standard travel time in the current implementation because of the service limitation. We would like to incorporate the departure time in future. Other future work includes more precise region determination because the travel time from each station is affected by road maps around the station.

ACKNOWLEDGMENTS

The authors would like to thank Val Laboratory Corporation for the generous support of Ekispert Web Service.

REFERENCES

- [1] A. MacEachren, F. Boscoe, D. Haug and L. Pickle. Geographic visualization: Designing manipulable maps for exploring temporally varying georeferenced statistics. *Proceedings IEEE Symposium on Information Visualization*, 87–94, 1998.
- [2] M. Migurski. Visualizing Urban Data. In Toby Segaran and Jeff Hammerbacher (Eds.) *Beautiful Data: The Stories Behind Elegant Data Solutions* Chapter 11, O'Reilly & Associates, 2009.
- [3] M. Mesbar, G. Currie, C. Lennon, and T. Northcott. Spatial and temporal visualization of transit operations performance data at a network level. *Journal of Transport Geography*, 25, 15–26, 2012.

- [4] S. Marie. Isochronous Application V0.8. <http://cartoo.dyndns.org/>
- [5] J. Soma. Triptrop NYC. <http://www.triptropnyc.com/>
- [6] S. Wehrmeyer. Mapnificent. <http://www.mapnificent.net/>
- [7] Google, Inc. Google Maps. <https://maps.google.com>
- [8] Google, Inc. Google Maps Utility Library v3. <https://code.google.com/p/google-maps-utility-library-v3/>
- [9] Khronos Group. WebGL - OpenGL ES 2.0 for the Web. <http://www.khronos.org/webgl/>
- [10] Vivid Solutions Inc. JTS Topology Suite. <http://www.vividsolutions.com/jts/JTSHome.htm>
- [11] Google, Inc. Google Geocoding API. <https://developers.google.com/maps/documentation/geocoding/>
- [12] Val Laboratory Corporation. Ekispert Web Service. <http://webservice.ekispert.com/>

ABOUT THE AUTHORS

1. Michio Shiraishi received the BA, MA, and PhD degrees from The University of Tokyo in 1997, 1999, 2003, respectively. He is a associate professor at Toho University.
2. Aya Inoue received the BSc from Toho University in 2014. She is currently a software engineer in U-DOM Co., Ltd.
3. Mikio Shinya is currently a Professor in the Department of Information Science, Toho University. He received a BSc in 1979, a MS in 1981, and a PhD in 1990 from Waseda University. He joined NTT Laboratories in 1981, and moved to Toho University in 2001.



Figure 5: The distribution of travel time from Tokyo Station (indicated as A). Although Shibuya (indicated as B) is closer Tokyo Station (A) than Kawasaki (C), users can easily find that the travel times from Tokyo station are similar.



Figure 6: The candidates of meeting point from Tokyo Station (indicated as A) and Shibuya Station(B). Since each station is colored according to the convenience of the station, we can find Yotsuya (C) and Ochanomizu (D) may be a good meeting point.

VISUALIZATION WITH VISIBILITY OF HIGHER DIMENSIONAL AND NON-EUCLIDEAN GEOMETRIES

János KATONA¹, Emil MOLNÁR², István PROK² and Jenő SZIRMAI²

¹Szent István University, Ybl Miklós Faculty of Architecture, Budapest, Hungary

²Budapest University of Technology and Economics, Department of Geometry, Budapest, Hungary

ABSTRACT: The theoretical background of our topic is the d -dimensional projective spherical space $\mathcal{PS}^d(\mathbf{V}^{d+1}; V_{d+1}; \mathbb{R}; \sim)$ or projective space \mathcal{P}^d , modelled as subspace incidence structure of the real $d + 1$ -dimensional vector space \mathbf{V}^{d+1} for points or its dual V_{d+1} for hyperplanes, respectively. Here \sim indicates the multiplicative equivalence by positive reals \mathbb{R}^+ in case \mathcal{PS}^d , or by non-zeros $\mathbb{R} \setminus \{0\}$ for \mathcal{P}^d . E.g. non-zero \mathbf{V}^{d+1} vectors $\mathbf{x} \sim c\mathbf{x}$ describe the same point $X(\mathbf{x})$ in \mathcal{PS}^d iff $c \in \mathbb{R}^+$. We report some new results in visualizing higher dimensional regular polytopes in Euclidean d -space, in particular in \mathbb{E}^4 (see e.g. [1], [2], [3] and [8]). Furthermore, we illustrate in \mathbb{E}^3 some new analogues of the classical objects; polyhedra, spheres, balls, their combinations in dense ball packing problems in the so-called Thurston 3-geometries $\mathbb{E}^3, \mathbb{S}^3, \mathbb{H}^3, \mathbb{S}^2 \times \mathbb{R}, \mathbb{H}^2 \times \mathbb{R}, \mathbf{SL}_2\mathbb{R}, \mathbf{Nil}$ and \mathbf{Sol} (see e.g. [4], [6] and [7]). To these last illustrations we have to specify the scalar product $\langle \cdot, \cdot \rangle$ in V_{d+1} so in \mathbf{V}^{d+1} ($d = 3$) by the signature and other requirements (as in [5], [6] and [7]).

Keywords: Higher-dimensional visualization, Thurston geometries, computer animation

1. INTRODUCTION

In our proceedings paper [2] we described the Coxeter-Schläfli theory of regular solids or d -polytopes, based on their characteristic simplices through Coxeter-Schläfli diagrams and matrices (see in Figure 1 and 2 next page).

1.1 Coxeter-Schläfli diagrams and matrices, illustrated by cubes and the 120-cell

We start with the usual 3-cube in the Euclidean 3-space \mathbb{E}^3 , where the barycentric simplex subdivision assigns the 3-dimensional centre A_3 , then A_2, A_1, A_0 as consecutive centres, a vertex A_0 is at the end (Fig. 1 and 2). We introduce the simplex faces: $b^0 = A_1A_2A_3$, opposite to vertex A_0 , and so on, $b^3 = A_0A_1A_2$, opposite to the cube centre A_3 . The square faces of the cube are characterized by the angle $\beta^{01} = \frac{\pi}{4} = \frac{\pi}{n_{01}}$ between the simplex faces b^0 and b^1 . The trigonal vertex figure of the 3-cube is described by the angle $\beta^{12} = \frac{\pi}{3} = \frac{\pi}{n_{12}}$ of faces b^1 and b^2 .

The "euclidicity" of 3-cube can be seen on the angle $\beta^{23} = \frac{\pi}{4} = \frac{\pi}{n_{23}}$ of faces b^2 and b^3 . Opposed to this, a spherical 3-cube in the spherical 3-space \mathbb{S}^3 has $\beta^{23} = \frac{\pi}{3} = \frac{\pi}{n_{23}}$ angle of b^2, b^3 , if it fills \mathbb{S}^3 with congruent copies. In the Bolyai-Lobachevskian hyperbolic 3-space \mathbb{H}^3 there is a space filler 3-cube with angle $\beta^{23} = \frac{\pi}{5} = \frac{\pi}{n_{23}}$ of b^2, b^3 , moreover, $\beta^{23} = \frac{\pi}{6} = \frac{\pi}{n_{23}}$ is also provides space filler cube in \mathbb{H}^3 , but with ideal vertices at infinity. 4-cube is illustrated in Fig. 3 and 4.

Here the basic facts are also indicated by the 120-cell (Fig. 5) and its matrix.

The extended Coxeter-Schläfli diagram $(5, 3, 3; \beta^{34} = \frac{2\pi}{5})$ with $\cos \beta^{34} = \frac{\sqrt{5}-1}{4}$ makes the determinant of the Coxeter-Schläfli matrix b^{ij} to be zero, according to the Euclidean signature $(+, +, +, +; 0)$ of the bilinear form $u_i b^{ij} v_j$ (or quadratic $u_i b^{ij} u_j$) by Einstein-Schouten convention. Thus we get the 120 regular 3-dodecahedra bounding the 4-polytope, now by parallel projection from the ideal line $E_3^\infty E_4^\infty$ as

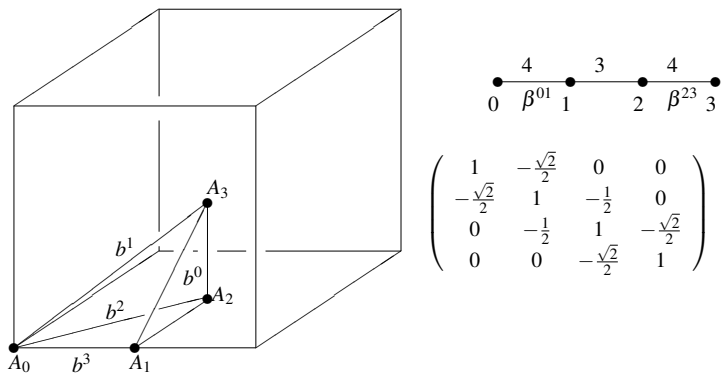
$$\{n_{01}, n_{12}, \dots, n_{d-2,d-1}; \beta^{d-1,d}\} \Longleftrightarrow$$

$$\begin{array}{ccccccc} & n_{01} & n_{12} & & & n_{d-2,d-1} & \\ \bullet & & \bullet & & \bullet & & \bullet \\ 0 & \beta^{01} & 1 & \beta^{12} & 2 & \dots & d-2 & d-1 & \beta^{d-1,d} & d \end{array}$$

$$b^{ij} = \begin{pmatrix} 1 & -\cos \beta^{01} & 0 & \dots & 0 & 0 \\ -\cos \beta^{01} & 1 & -\cos \beta^{12} & \dots & 0 & 0 \\ 0 & -\cos \beta^{12} & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & -\cos \beta^{d-1,d} \\ 0 & 0 & 0 & \dots & -\cos \beta^{d-1,d} & 1 \end{pmatrix}$$

where $\beta^{ij} = \frac{\pi}{n_{ij}}$ for $i, j = 0, 1, \dots, d; i \neq j, (i, j) \neq (d-1, d); 1 < n_{ij} \in \mathbb{N}$ natural numbers.

Figure 1: Coxeter-Schläfli-diagram and matrix



$$\begin{array}{ccccccc} & 4 & & 3 & & 3 & & 4 & \\ \bullet & & \bullet & & \bullet & & \bullet & & \bullet \\ 0 & \beta^{01} & 1 & & 2 & & 3 & & \beta^{34} & 4 \end{array}$$

$$\begin{pmatrix} 1 & -\frac{\sqrt{2}}{2} & 0 & 0 & 0 \\ -\frac{\sqrt{2}}{2} & 1 & -\frac{1}{2} & 0 & 0 \\ 0 & -\frac{1}{2} & 1 & -\frac{1}{2} & 0 \\ 0 & 0 & -\frac{1}{2} & 1 & -\frac{\sqrt{2}}{2} \\ 0 & 0 & 0 & -\frac{\sqrt{2}}{2} & 1 \end{pmatrix}$$

Figure 2: Cube in $\mathbb{E}^3, \mathbb{E}^4$ and symbols for them

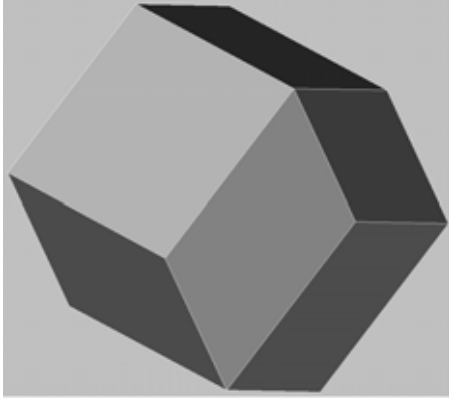


Figure 3: The 4-cube (8-cell) with Coxeter-Schläfli symbol $(4, 3, 3; 4)$

centre figure $C[C_3(\mathbf{e}_3 = \mathbf{e}_{p+1} = \mathbf{e}_3), C_d(\mathbf{e}_d = \mathbf{e}_4)]$ into the picture plane $\Pi(P_0 P_1^\infty P_2^\infty)$. The $p = 2$ -dimensional picture can have (and seems to have now) more contour components (at the animation [8] near black parts where new $p = 2$ -faces appear or disappear with flashy scattering), the white pentagons are almost parallel to the picture plane.

Imagine also 120-cells $(5, 3, 3; 3)$, $(5, 3, 3; 4)$, $(5, 3, 3; 5)$, each of them (with its congruent copies) fills the Bolyai-Lobachevskian hyperbolic space \mathbb{H}^4 , with proper vertices. All three have the signature $(+, +, +, +; -)$.

2. ANIMATION BY REGULAR 4-POLYTOPES IN ORTHOGONAL OR CENTRAL PROJECTION INTO THE COMPUTER 2-SCREEN

Our principles have been described in [1], [2] and [3], now the parallel projection is visualised in the homepage [8]. for free download. Here our further developments and new polytopes will be animated, as well. We summarize only the most important steps, using the machinery of the projective geometry. That is of the $d = 4$ -dimensional projective spheri-

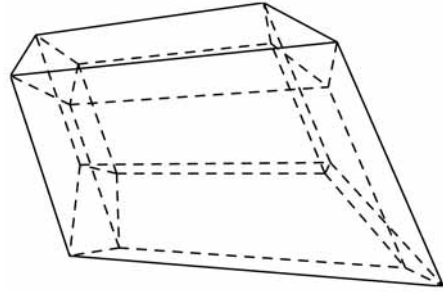


Figure 4: Central projection of a four-dimensional cube

cal space $\mathcal{P}\mathcal{S}^d(\mathbf{V}^{d+1}; \mathbf{V}_{d+1}; \mathbb{R}; \sim)$ or projective space \mathcal{P}^d . Both are modelled as subspace incidence structure of the real $d + 1$ -dimensional vector space \mathbf{V}^{d+1} for points, or its dual \mathbf{V}_{d+1} for hyperplanes, respectively. Here \sim indicates the multiplicative equivalence by positive reals \mathbb{R}^+ in case $\mathcal{P}\mathcal{S}^d$, or by non-zeros $\mathbb{R} \setminus \{0\}$ for \mathcal{P}^d . E.g. non-zero \mathbf{V}^{d+1} vectors $\mathbf{x} \sim c\mathbf{x}$ describe the same point $X(\mathbf{x})$ in $\mathcal{P}\mathcal{S}^d$ iff $c \in \mathbb{R}^+$.

A hyperplane $u(\mathbf{u})$ represented by a non zero form $\mathbf{u} \in \mathbf{V}_{d+1}$, is incident to a point $X(\mathbf{x})$, as above, iff $\mathbf{x}\mathbf{u} = 0$. $X(\mathbf{x})$ lies in the positive half-sphere of $u(\mathbf{u})$, iff $\mathbf{x}\mathbf{u} > 0$, etc.

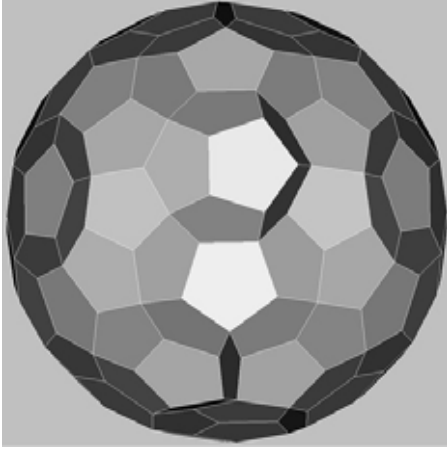
Our projective sphere $\mathcal{P}\mathcal{S}^d$ and our regular d -polytope \mathcal{P} will be considered by a Cartesian projective world coordinate simplex (WCS) by a dual basis pair $\{\mathbf{e}_0; \mathbf{e}_1, \dots, \mathbf{e}_d\} \subset \mathbf{V}^{d+1}$ and $\{\mathbf{e}^0; \mathbf{e}^1, \dots, \mathbf{e}^d\} \subset \mathbf{V}_{d+1}$ with $\mathbf{e}_i \mathbf{e}^j = \delta_i^j$ the Kronecker symbol.

The proper origin $E_0(\mathbf{e}_0)$ will be combined by the ideal (infinite) points $E_i^\infty(\mathbf{e}_i)$ of the coordinate i -axes, $i \in \{1, \dots, d\}$. So our WCS will be

$$\{E_0(\mathbf{e}_0); E_1(\mathbf{e}_1), \dots, E_d(\mathbf{e}_d); E(\mathbf{e})\}$$

$$\mathbf{e} = \mathbf{e}_0 + \mathbf{e}_1 + \dots + \mathbf{e}_d$$

where the above unit point $E(\mathbf{e})$ normalizes the bases with a Euclidean scalar product $\langle ; \rangle$ of signature $(+, \dots, +; 0)$ as usual for the Euclidean projective modelling later on. The visualization



$$\begin{pmatrix} 1 & -\frac{\sqrt{5}+1}{4} & 0 & 0 & 0 \\ -\frac{\sqrt{5}+1}{4} & 1 & \frac{-1}{2} & 0 & 0 \\ 0 & \frac{-1}{2} & 1 & \frac{-1}{2} & 0 \\ 0 & 0 & \frac{-1}{2} & 1 & -\frac{\sqrt{5}-1}{4} \\ 0 & 0 & 0 & -\frac{\sqrt{5}-1}{4} & 1 \end{pmatrix}$$

Figure 5: The 120-cell in \mathbb{E}^4 with Coxeter-Schläfli symbol $(5,3,3) \beta^{34} = \frac{2\pi}{5}$ and its matrix

will be summarised in the following algorithmic steps (as also sketched in [2] Sect. 2-5).

1. Our $d = 4$ -polytope \mathcal{P} is not necessarily regular but it shall have (say for simplicity) a *barycentric simplex subdivision*. The (formal) d -centre of \mathcal{P} will be denoted by A_d , then comes a $(d-1)$ -centre of a $(d-1)$ -hyperface of \mathcal{P} , ..., a $p = 2$ -centre $A_{p=2}$ of its consecutive $p = 2$ -face, an edge 1-centre A_1 , a 0-vertex A_0 . The so obtained barycentric simplices have *flag-lattice structure* by so-called *k-adjacencies* $k \in \{0, \dots, d-1\}$.

Namely, each barycentric simplex

$$D = A_0 \dots A_k \dots A_d = b^0 \dots b^k \dots b^d$$

has a k -hypersimplex

$$b^k = A_0 \dots A_{k-1} A_{k+1} \dots A_d$$

along which there is exactly one k -adjacent barycentric simplex of \mathcal{P} denoted by $\sigma^k D$. An occasional symmetry operation γ will be written

on the right into exponent as $(\sigma^k D)^\gamma = \sigma^k (D^\gamma)$ indicates this formal associativity law. We do not cite more details of the "D-symbol theory" [5]. Let us only mention that the adjacency structure of the finitely many symmetry orbits of barycentric simplices and the symmetric matrix function $\mathcal{M} = m^{ij}(D)$, relying on $\sigma^j \sigma^i$ operations; combinatorially describe our polytope \mathcal{P} ; can canonically enumerate its barycentric simplices; the elements of its finite symmetry group, generated by adjacency reflections; etc.

Let us remark that our regular $d = 4$ -polytope \mathcal{P} can be given by its unique characteristic simplex and its Coxeter-Schläfli symbol [2].

2. We assume that our \mathcal{P} is given in our WCS = $\{E_0; E_1; \dots; E_{d=4}; E\}$ by the vertices of all barycentric simplices with coordinate $d+1$ -tuples. For $p = 2$ -projection we use only some A_0 -, A_1 -, $A_{p=2}$ -type centres to our later visibility procedure. The centre A_d of \mathcal{P} is fixed into the origin E_0 . \mathcal{P} is standing, while the observer as a Camera $\{\Pi, C\}$ will move by the time t in a

prescribed way (as attractive as possible !?).

3. Let the picture plane (computer screen)

$$\Pi[P_0(\mathbf{p}_0); P_1^\infty(\mathbf{p}_1); \dots; P_{p=2}^\infty(\mathbf{p}_{p=2})]$$

and the complementary centre figure

$$C[C_{p+1}(\mathbf{c}_{p+1} = \mathbf{c}_3); \dots; C_{d=4}^\infty(\mathbf{c}_d = \mathbf{c}_4)]$$

be given as a new coordinate simplex by the matrix scheme see Figure 6.

4. Any point $X(\mathbf{x})$ in the visible region can be expressed as

$$\mathbf{x} \sim (1; x^1, \dots, x^{p=2}, x^{p+1}, \dots, x^d) \begin{pmatrix} \mathbf{e}_0 \\ \vdots \\ \mathbf{e}_d^\infty \end{pmatrix} \sim (y^0, y^1, \dots, y^{p=2}, y^{p+1}, \dots, y^d) (Cam) \begin{pmatrix} \mathbf{e}_0 \\ \vdots \\ \mathbf{e}_d^\infty \end{pmatrix}$$

so that for $y^0 > 0$ holds

$$(1; x^1, \dots, x^{p=2}, x^{p+1}, \dots, x^d) (Cam)^{-1} \sim (1, \frac{y^1}{y^0}, \dots, \frac{y^{p=2}}{y^0}, \frac{y^{p+1}}{y^0}, \dots, \frac{y^d}{y^0}).$$

The above equation shows

$$\mathbf{x} \sim \mathbf{y} + \mathbf{c} \sim y^0 \mathbf{p}_0 + y^1 \mathbf{p}_1 + \dots + y^{p=2} \mathbf{p}_{p=2} + c^{p+1} \mathbf{c}_{p+1} + \dots + c^d \mathbf{c}_d,$$

where $Y(\mathbf{y}) := {}^p X(\mathbf{x})$ is the projection of $X(\mathbf{x})$ into Π .

5. The relative visibility of $X(\mathbf{x})$ to $X'(\mathbf{x}')$ can be defined by the following convention below:

a) if the images $({}^p \mathbf{x}) \sim (\mathbf{y})$ and $({}^p \mathbf{x}') \sim (\mathbf{y}')$ are different, both are visible;

b) if the images are the same, i.e. $(\mathbf{y}) \sim (\mathbf{y}')$, then $\frac{c^{p+1=3}}{y^0} \geq \frac{c^{p+1=3'}}{y^{0'}}$, ...

c) if above equalities hold, then $\frac{c^{d=4}}{y^0} \leq \frac{c^{d'=4'}}{y^{0'}}$, the reverse inequality holds (see also our zero

convention in step 10.). Then $X(\mathbf{x})$ is closer to the centre figure C than $X'(\mathbf{x}')$. The above conventions naturally extend the $d = 3$ -dimensional visibility (see Figure 7).

6. The above relative visibility will be extended to the case where a point lies over (say) a $p = 2$ -face f of \mathcal{P} , namely, then $X(\mathbf{x})$ cover its projection $Z(\mathbf{z})$ into the above $p = 2$ -face f of \mathcal{P} by the above projection with $[f, C]$ instead of $[\Pi, C]$.

7. However, the above procedure would be too complicated to a real time animation, so we specify it in the following only to the contour components of \mathcal{P} by the time t under a given projection procedure below.

Moreover, we define a conventional lightening procedure, instead of putting a light source into the centre point $C_{p+1=3}$ of the centre figure $C[C_{p+1}, \dots, C_d]$. Namely, the brightness (by shading) of a $p = 2$ -face ${}^p A_0 {}^p A_1 {}^p A_{p=2}$ will be proportional with its (relative) $p = 2$ -area.

8. The moving (by $t \geq 0$) observer (Cam) will be specified by

$$P_0[\mathbf{p}_0 = 1 \cdot \mathbf{e}_0 + \rho(t) \cos(\omega_1 t) \mathbf{e}_1 + \rho(t) \sin(\omega_1 t) \mathbf{e}_2]$$

$$P_1[\mathbf{p}_1 = -\cos(\omega_1 t) \mathbf{e}_1 - \sin(\omega_1 t) \mathbf{e}_2]$$

$$P_2[\mathbf{p}_2 = \sin(\omega_1 t) \mathbf{e}_1 - \cos(\omega_1 t) \mathbf{e}_2], \dots, P_p[\mathbf{p}_p = \mathbf{e}_p]$$

for $p > 2$ dimensional screen, if occurs.

$$C_{p+1=3}[\mathbf{c}_{p+1} = 1 \cdot \mathbf{e}_0 + \rho(t) \cos(\omega_1 t) \mathbf{e}_1 + \rho(t) \sin(\omega_1 t) \mathbf{e}_2 + h(t) \mathbf{e}_{p+1}, \dots, C_{d=4}[\mathbf{c}_d = \mathbf{e}_d]$$

The given distance function

$$\rho(t) = \rho_0 + (\rho_1 - \rho_0) \sin(\omega_2 t)$$

can be periodic or constant. Similarly, the height function

$$h(t) = h_0 + (h_1 - h_0) \sin(\omega_3 t)$$

can be prescribed accordingly.

Think of $d = 4$, $p = 2$, $s = d - p - 1 = 1$ (Figure 7). The picture plane Π is fixed to the computer screen. P_0 is in distance $\rho(t)$ from the origin E_0 of WCS. Imagine the eye of the observer

$$\begin{pmatrix} \mathbf{p}_0 \\ \vdots \\ \mathbf{p}_{p=2} \\ \mathbf{c}_{p+1} \\ \vdots \\ \mathbf{c}_{d=4} \end{pmatrix} = \begin{pmatrix} 1 & \dots & p_0^p & p_0^{p+1} & \dots & p_0^d \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & p_p^p & p_p^{p+1} & \dots & p_p^d \\ c_{p+1}^0 & \dots & c_{p+1}^p & c_{p+1}^{p+1} & \dots & c_{p+1}^d \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & c_d^p & c_d^{p+1} & \dots & c_d^d \end{pmatrix} \begin{pmatrix} \mathbf{e}_0 \\ \vdots \\ \mathbf{e}_{p=2} \\ \mathbf{e}_{p+1} \\ \vdots \\ \mathbf{e}_{d=4} \end{pmatrix} = (\text{Cam}) \begin{pmatrix} \mathbf{e}_0 \\ \vdots \\ \mathbf{e}_d \end{pmatrix}$$

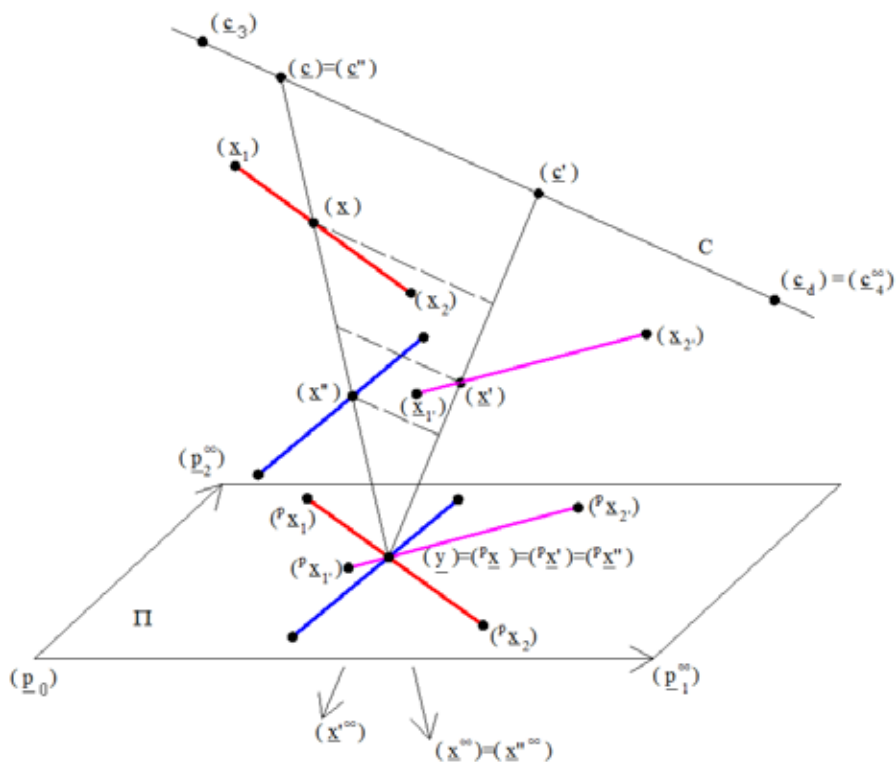
Figure 6: *Definition of Camera*

Figure 7: *Projection of segments to extended local visibility*

in $C_{p+1=3}$ in height $h(t)$. The observer (Cam) is rotating about \mathcal{P} , and changing the parameters.

9. The inverse matrix $(\text{Cam})^{-1}$ will be simple now (by expressing the basis $\{\mathbf{e}_i\}$ by $\{\mathbf{p}_i; \mathbf{c}_j\}$ as usual). Thus we obtain the picture

$$\begin{aligned} {}^pX^1 &= \frac{y^1}{y^0} = \\ &= \frac{h(t)}{h(t) - x^{p+1}} [\rho(t) - x^1 \cos(\omega_1 t) - x^2 \sin(\omega_1 t)] \\ {}^pX^2 &= \frac{y^2}{y^0} = \\ &= \frac{h(t)}{h(t) - x^{p+1}} [x^1 \sin(\omega_1 t) - x^2 \cos(\omega_1 t)], \dots, \\ \dots, {}^pX^p &= \frac{y^p}{y^0} = \frac{h(t)x^p}{h(t) - x^{p+1}} \end{aligned}$$

and furthermore

$$\frac{c^{p+1=3}}{y^0} = \frac{x^{p+1}}{h(t) - x^{p+1}}, \dots, \frac{c^{d=4}}{y^0} = \frac{h(t)x^d}{h(t) - x^{p+1}}$$

As above in step 5, these last coordinates order the closeness of 0-vertices A_0 , edge 1-centres A_1 , face $p = 2$ -centre A_2 of \mathcal{P} to the centre figure C (as Figure 7 indicates).

10. The visibility and brightness of the $p = 2$ -faces of \mathcal{P} from the centre figure C above, at any moment $t \geq 0$, will be decided by determining the visible contour components. To these we calculate only some $p + 1 = 3$ order image determinants, the area ${}^pA_0{}^pA_1{}^pA_2$, more precisely its sign or vanishing (up to a given small positive ε).

The (or a) highest $p = 2$ -face, with the (a) largest $\frac{c^{p+1=3}}{y^0}$ coordinate of its $p = 2$ -centre A_2 , is surely visible. Then we go further to its $p = 2$ adjacent $p = 2$ -face $A'_0A'_1A'_2$ with $A_0A_1 = A'_0A'_1$, $A'_2 \neq A_2$.

a) If $\det({}^p\mathbf{a}'_0, {}^p\mathbf{a}'_1, {}^p\mathbf{a}'_2)$ change sign related to $\det({}^p\mathbf{a}_0, {}^p\mathbf{a}_1, {}^p\mathbf{a}_2)$ then both $p = 2$ -faces are visible, with brightness of their area, respectively.

b) If above the sign preserves, then $A'_0A'_1 = A_0A_1$ is a contour edge, ${}^pA'_0{}^pA'_1{}^pA'_2$ is not visible.

c) If $\det({}^p\mathbf{a}'_0, {}^p\mathbf{a}'_1, {}^p\mathbf{a}'_2) = 0$, then $A'_0A'_1A'_2$ is a contour $p = 2$ -face. In such a way, maybe experimentally (no more detailed), we take the visible

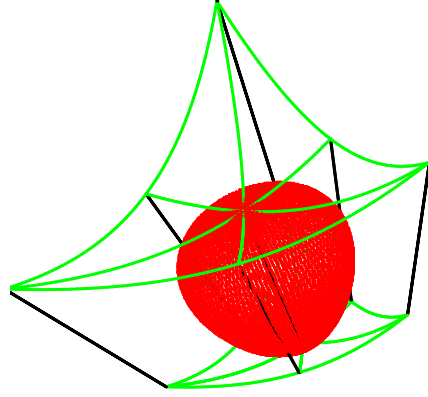


Figure 8: Sphere (ball) in a prism of $\mathbb{H}^2 \times \mathbb{R}$ to its cone model

$p = 2$ -faces by the order of $p = 2$ -centres A_2 related to the centre figure C .

3. PICTURES TO THURSTON 3-SPACES

The projective interpretations of the 8 Thurston geometries $\mathbb{E}^3, \mathbb{S}^3, \mathbb{H}^3, \mathbb{S}^2 \times \mathbb{R}, \mathbb{H}^2 \times \mathbb{R}, \mathbf{SL}_2\mathbb{R}, \mathbf{Nil}$ and \mathbf{Sol} in our former papers (see e.g. in [4], [5], [6] and [7]) in \mathcal{PS}^3 and in \mathcal{P}^3 give us the possibility to model them in the Euclidean space as well.

Here we present only some attractive pictures as preindicated in our abstract: Figure 8, 9, 10, 11, 12, 13, 14, 15 and 16.

REFERENCES

- [1] J. Katona and E. Molnár. Visibility of the higher-dimensional central projection into projective sphere. *Acta Mathematica Hungarica*, 123(3): 291–309, 2009.
- [2] J. Katona, E. Molnár, and I. Prok. Visibility of the 4-dimensional regular solids, moving on the computer screen. In *Proceedings of the 13th International Conference on Geom-*

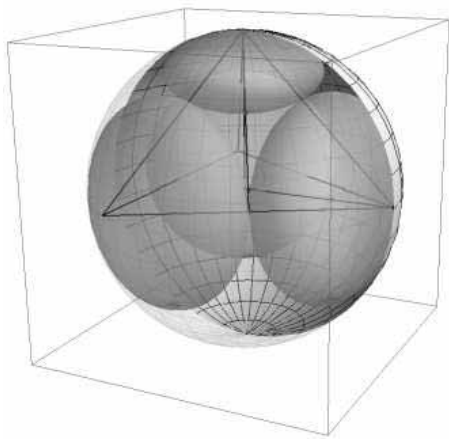


Figure 9: A densest horoball packing to the Coxeter group $(3, 3, 6)$ to the regular ideal simplex tiling in the hyperbolic space \mathbb{H}^3 , in its Beltrami-Klein model

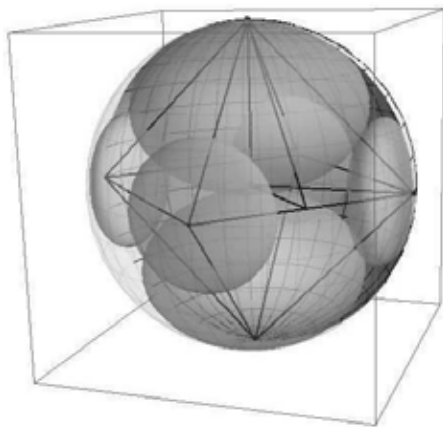


Figure 11: An analogous (as before) horoball packing to $(3, 4, 4)$ of the regular ideal octahedron tiling in \mathbb{H}^3

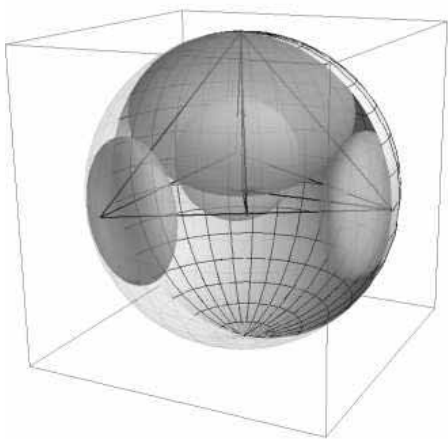


Figure 10: Another horoball packing in \mathbb{H}^3 with the same density

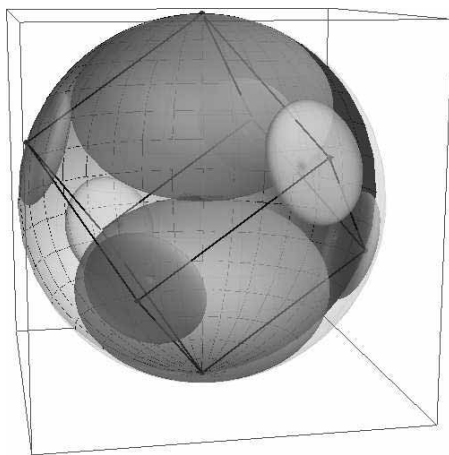


Figure 12: Another analogue to $(4, 3, 6)$ of the regular ideal cube tiling in \mathbb{H}^3

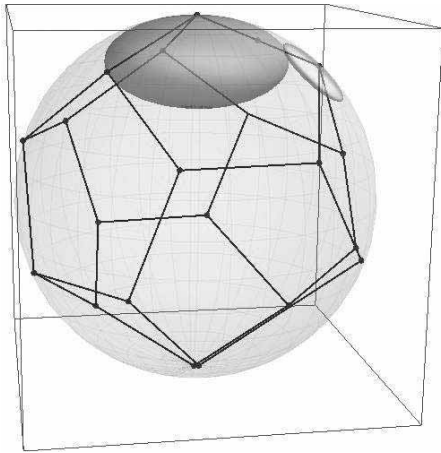


Figure 13: Another analogue to $(5, 3, 6)$ of the regular ideal dodecahedron tiling in \mathbb{H}^3

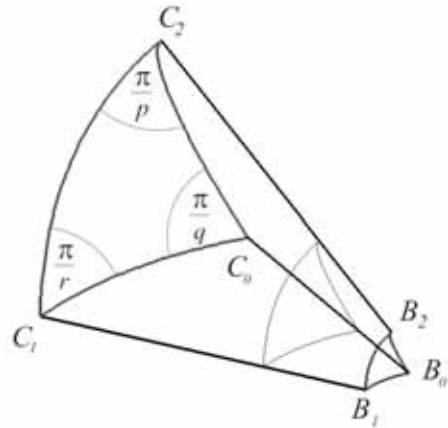


Figure 15: A prism D-V cell in $\mathbb{S}^2 \times \mathbb{R}$

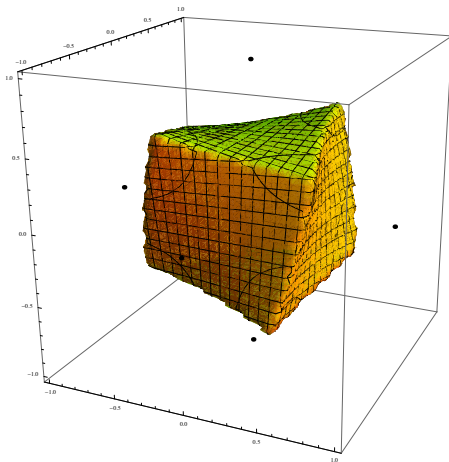


Figure 14: A Nil-prism as a Dirichlet-Voronoi cell to a geodesic (or translation) ball packing (by János Pallagi and Benedek Schultz)

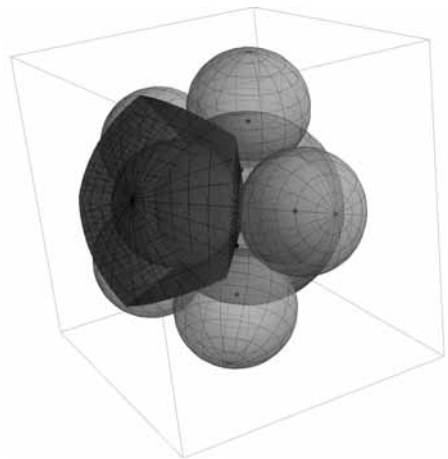


Figure 16: A ball packing in $\mathbb{S}^2 \times \mathbb{R}$

etry and Graphics, Dresden, Germany. TUD, 2008.

- [3] J. Katona, E. Molnár, I. Prok, and J. Szirmai. Higher-dimensional central projection into 2-plane with visibility and application. *Kragujevac Journal of Mathematics*, 35(2): 249–263, 2011.
- [4] E. Molnár. The projective interpretation of the eight 3-dimensional homogeneous geometries. *Beiträge zur Algebra und Geometrie*, 38(2): 261–288, 1997.
- [5] E. Molnár. On d-symbols and orbifolds in an algorithmic way. *Atti Semin. Mat. Fis. Univ. Modena*, 58: 263–276, 2011.
- [6] E. Molnár and J. Szirmai. Symmetries in the 8 homogeneous 3-geometries. *Symmetry: Culture and Science*, 21(1–3): 87–117, 2010.
- [7] E. Molnár, J. Szirmai, and A. Vensin. Packings by translation balls in $\mathbf{SL}_2\mathbb{R}$. *To appear in J. Geometry*, 2014.
- [8] I. Prok. <http://www.math.bme.hu/~prok/NCells/index.html>. (*Online, accessed 10.06.2014*).
- 4. Jenő SZIRMAI, PhD, docent of BME Department of Geometry. H-1111 Budapest Egry József u. 1. H. II. 22., Hungary. E-mail: szirmai@math.bme.hu, URL: www.math.bme.hu/~szirmai.

ABOUT THE AUTHORS

- 1. János KATONA, PhD, lecturer of SZIE Ybl Miklós Faculty of Architecture. Address: H-1146 Budapest Thököly u. 74., Hungary. E-mail: katona.janos@ybl.szie.hu, URL: www.katonajanos.hu.
- 2. Emil MOLNÁR, Dr. habil. prof. emer. of BME Department of Geometry. Address: H-1111 Budapest Egry József u. 1. H. II. 22., Hungary. E-mail: emolnar@math.bme.hu, URL: www.math.bme.hu/~emolnar.
- 3. István PROK, PhD, docent of BME Department of Geometry. H-1111 Budapest Egry József u. 1. H. II. 22., Hungary. E-mail: prok@math.bme.hu, URL: www.math.bme.hu/~prok.

VISUALIZING NORMAL EQUATIONS IN LEAST-SQUARES ADJUSTMENT

Andreas RONCAT

Research Groups Photogrammetry and Remote Sensing,
Department of Geodesy and Geoinformation, Vienna University of Technology, Austria

ABSTRACT: The adjustment by least squares dates back to more than 200 years—commonly attributed to C.F. Gauss [1]—and is widely applied in practically all disciplines where linear and non-linear regression is sought. These disciplines include natural sciences, engineering and also social sciences. German geodesist W. Niemeier entitled the adjustment by least squares as a *brittle beauty* („spröde Schöne“ in the original German text [6]), highlighting both the possible rejection of a scholar at first sight and the attraction for a scientist dealing closer with this topic.

Besides its stochastic content, the setup of the least-squares adjustment approach by solving normal equations is not at last a geometric task. However, this aspect is seldom given great attention in relevant literature, especially in introductory textbooks on this topic.

However, special configurations allow even an examination from the viewpoint of descriptive geometry: This is the case if the parameter space does not exceed four dimensions and the space of observations can be embedded into it.

In order to reduce the *brittle* aspect mentioned above, this paper is intended to highlight the geometry of least-squares adjustment by visualizing it at the most possible level. It presents examples for the setup and solutions of such least-squares adjustment problems in two and three dimensions. The case of four-dimensional adjustments is discussed as well.

Keywords: Least-Squares Adjustment, Descriptive Geometry, Normal Equations

1. INTROCUCTION

The adjustment by least squares has seen a 200-years-long history and is in use in a wide range of disciplines. Traditionally, this method is introduced in specific courses and textbooks by means of its stochastic properties, primarily because it yields a *best fit* in case of observations containing random errors only, i.e. observations without systematic errors. However, the actual adjustment step by solving a normal equation system can also be seen as a geometric task.

In this paper, we want to derive and present the least-squares adjustment in a purely (Euclidean) geometric sense yielding a *closest* fit to the observations. This can be done without introducing the stochastic properties of the least-squares adjustment, which are given in Section 5.

For the geometric derivation in low dimensions, we use techniques of descriptive geometry. In four or less dimensions, the solution of the normal equation can be visualized as intersection of hyperplanes; the link to the observations is possible if these observations can be embedded into the parameter space.

2. DERIVATION BY MEANS OF A TWO-DIMENSIONAL EXAMPLE

We will start the presentation with the example of a parabola in the xy plane. As this curve is defined as the the locus of the points with equal distance to both a point $\mathbf{f} = (f_x, f_y)^\top$ and to a line l , this gives 4 parameters for a unique characterization of a parabola.

Let in this example the line l (directrix) be the x

axis, i.e. $y = 0$. Thus, the determination of $u = 2$ coordinates of the point \mathbf{f} (focus) is sufficient to uniquely determine such a parabola. Its equation now reads as follows:

$$y = \sqrt{(x - f_x)^2 + (y - f_y)^2},$$

or

$$y - \frac{1}{2f_y}(x - f_x)^2 - \frac{1}{2}f_y = 0. \quad (1)$$

Given 3 or more points with observed coordinates x_i and y_i , we want to find the parabola in the mentioned configuration being “closest” to the observed values. This x and y coordinate are given in the same coordinate system as the parameters f_x and f_y .

Assumed that we have already determined approximate values $f_{x,0}$ and $f_{y,0}$ of the unknown parameters, Equation (1) reads as follows:

$$\varphi_i \dots y_i - \frac{1}{2f_{y,0}}(x_i - f_{x,0})^2 - \frac{1}{2}f_{y,0} = w_i. \quad (2)$$

This kind of equations, also referred to as *condition equation* [5], may result in non-zero values w_i meaning that the approximate set of parameters is not fully compatible with the observed values. These w_i are summarized to a column vector \mathbf{w} (discrepancy vector). Its dimension is determined by the number n of observed points. Please note that at this point, the observed values x_i and y_i are *not* treated as stochastic but as deterministic quantities.

In order to visualize the geometric situation, we will investigate the example of $n = 3$ observed point so that the problem is overdetermined (see Figure 1).

The aim for a closest fit motivates the minimization of the norm of the discrepancy vector \mathbf{w} :

$$\|\mathbf{w}\| \rightarrow \min. \Leftrightarrow \|\mathbf{w}\|^2 = \mathbf{w}^T \mathbf{w} = \sum_{i=1}^3 w_i^2 \rightarrow \min.$$

which explains the choice of the term “least squares”.

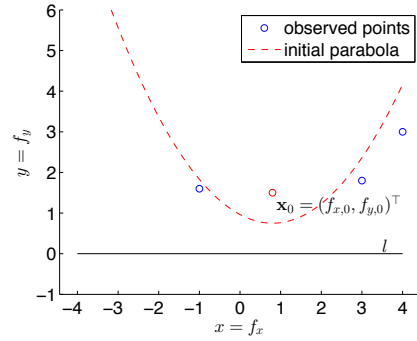


Figure 1: Observed points $\mathbf{p}_i = (x_i, y_i)^T$ and approximate solution for a parabola (red). In this example, the coordinate differences in y direction of the observed points to the parabola correspond to the discrepancies w_i .

The discrepancy vector \mathbf{w} is considered as $\mathbf{w}(f_x, f_y)$. A linear Taylor approximation around $(f_{x,0}, f_{y,0})$ gives

$$\begin{aligned} \mathbf{w}(f_{x,0} + \Delta f_x, f_{y,0} + \Delta f_y) &= \mathbf{w}(\mathbf{x}_0 + \Delta \mathbf{x}) \\ &\simeq \mathbf{w}(\mathbf{x}_0) + \mathbf{J} \Delta \mathbf{x}, \end{aligned} \quad (3)$$

with the $(n \times u)$ matrix \mathbf{J} being the Jacobian matrix w.r.t. the unknowns. In our example, it is a (3×2) matrix whose entries are given by the derivatives

$$\frac{\partial \varphi_i}{\partial f_x} = \frac{1}{f_y}(x - f_x) \text{ and } \frac{\partial \varphi_i}{\partial f_y} = \frac{1}{2f_y^2}(x - f_x)^2 - \frac{1}{2}.$$

With the $n = 3$ discrepancies and the 2 parameters, $\mathbf{w}(f_x, f_y)$ can be interpreted as a 2-dimensional surface in 3-space, as displayed in Figure 2.

A necessary condition for minimization is the first derivative resulting to zero:

$$\frac{d\mathbf{w}^T \mathbf{w}}{d\mathbf{x}} = 0,$$

therefore Equation (3) results in

$$2d\Delta \mathbf{x} \mathbf{J}^T \mathbf{w}(\mathbf{x}_0) + 2d\Delta \mathbf{x} \mathbf{J}^T \mathbf{J} \Delta \mathbf{x} = 0,$$

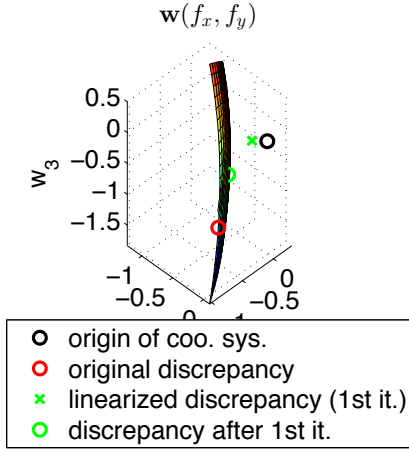


Figure 2: Discrepancy vector \mathbf{w} as 2-parameter surface in 3-space. The resulting discrepancy for the approximate values \mathbf{x}_0 is indicated by the red circle; within the first iteration of the adjustment, the green circle has been reached via linearization; the point indicated by the green cross is the origin's foot of perpendicular in the tangent plane at the red circle.

or

$$\underbrace{\mathbf{J}^T \mathbf{J}}_{=: \mathbf{N}} \Delta \mathbf{x} = -\mathbf{J}^T \mathbf{w}(\mathbf{x}_0).$$

The matrix $\mathbf{N} = \mathbf{J}^T \mathbf{J}$ is known as *normal equation matrix* [5], as its rows contain the normal vectors of hyperplanes in the parameter space, i.e. the normal vectors of lines in the current example:

$$\underbrace{\begin{pmatrix} \mathbf{n}_1^T \\ \mathbf{n}_2^T \end{pmatrix}}_{=: \mathbf{N}} \begin{pmatrix} \Delta f_x \\ \Delta f_y \end{pmatrix} = -\begin{pmatrix} \mathbf{j}_1^T \mathbf{w} \\ \mathbf{j}_2^T \mathbf{w} \end{pmatrix}$$

where \mathbf{j}_i are the column vectors of the Jacobian \mathbf{J} .

We see that the solution of this *normal equations* is just the intersection of n hyperplanes in the parameter space. The solution is unique for a regular normal equation matrix \mathbf{N} . The solution for the example of a parabola is shown in Figure 3.

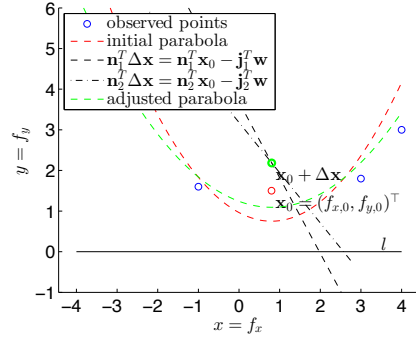


Figure 3: First iteration of the least-squares solution for the parabola fit of Figure 1. The updated focus point (green circle) is found as the intersection of the black dotted line and the black dash-dotted line, each representing one line of the normal equation system.

For $\mathbf{x}_0 + \Delta \mathbf{x}$, a new value for \mathbf{w} is retrieved (see Figure 2) as well as a new value for \mathbf{J} . The adjustment procedure is repeated iteratively until convergence, i.e. until the update in the parameter vector $\Delta \mathbf{x}$ does not cause a significant change in the discrepancy vector.

3. FURTHER EXAMPLES IN TWO AND THREE DIMENSIONS

In this section, we aim at the least-squares parameter estimation of a circle with (a) known and (b) unknown radius. While the first example is also 2-dimensional, the latter is a 3-dimensional task.

The equation of a circle in a plane,

$$(x - m_x)^2 + (y - m_y)^2 = r^2,$$

already gives us the condition equation for our first example:

$$\varphi_i \dots (x_i - m_{x,0})^2 + (y_i - m_{y,0})^2 - r^2 = w_i. \quad (4)$$

With the derivatives w.r.t. the unknowns m_x and m_y ,

$$\frac{\partial \varphi_i}{\partial m_x} = -2(x - m_x) \text{ and } \frac{\partial \varphi_i}{\partial m_y} = -2(y - m_y),$$

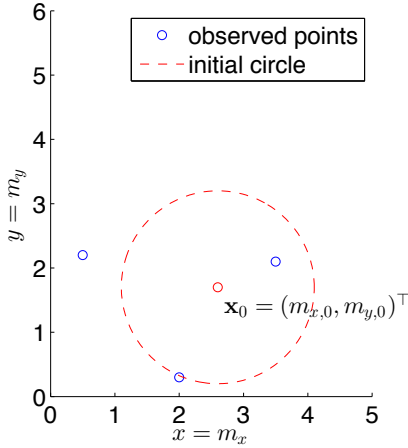


Figure 4: Observed points $\mathbf{p}_i = (x_i, y_i)^\top$ and approximate solution for a circle (red).

the Jacobian matrix \mathbf{J} and the normal equation matrix \mathbf{N} can be set up as above; the solution of the normal equation system can be interpreted as the intersection of two straight lines in the 2D parameter space as well; see Figures 4 and 6.

Again, 3 points with coordinates x_i and y_i were observed so that the discrepancy vector $\mathbf{w}(m_x, m_y)$ can also be visualized as 2D surface in 3D space, as shown in Figure 5.

In the next example, we will add the radius r as third unknown to the adjustment problem, with r_0 as initial estimate for this parameter. Thus, the condition equations change slightly to

$$\varphi_i \dots (x_i - m_{x,0})^2 + (y_i - m_{y,0})^2 - r_0^2 = w_i. \quad (5)$$

and the Jacobian matrix \mathbf{J} changes its size to $(n \times 3)$, with

$$\frac{\partial \varphi_i}{\partial r} = -2r_0$$

as entries of its third column. As a consequence, the normal equation matrix \mathbf{N} has a size of (3×3) . The geometric equivalent to the solution of the normal equation system is the intersection of three planes in 3D space.

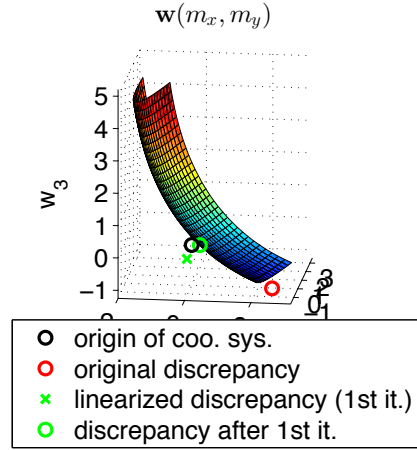


Figure 5: Discrepancy vector $\mathbf{w}(m_x, m_y)$ of a circle as 2-parameter surface in 3-space.

We consider a set of $n = 11$ observed points in this example (see Figure 7), yielding an (11×1) discrepancy vector \mathbf{w} . Therefore, it cannot be visualized as a surface in 3D space as in the examples before.

Figures 8 and 9 illustrate the geometric content of the solution for the normal equation system. We will use a technique known from descriptive geometry to retrieve the solution [8]: The three planes given by the equations

$$\mathbf{n}_i^\top \Delta \mathbf{x} = \mathbf{n}_i^\top \mathbf{x}_0 - \mathbf{j}_i^\top \mathbf{w}(\mathbf{x}_0)$$

are represented by their trace triangles in the planes parallel to the coordinate planes at \mathbf{x}_0 ; they are indicated by black solid lines, black dotted lines and black dash-dotted lines, resp. The intersections of corresponding traces and the subsequent connection of two intersection points give the intersection lines between two of the planes, illustrated by the dotted cyan, magenta and yellow lines, resp. Given that none of these lines is parallel to another, these three lines have of course *one* common intersection point, which is the solution of the normal equation system, i.e. $\mathbf{x}_0 + \Delta \mathbf{x}$ —or $\Delta \mathbf{x}$ in the coordinates reduced to \mathbf{x}_0 .

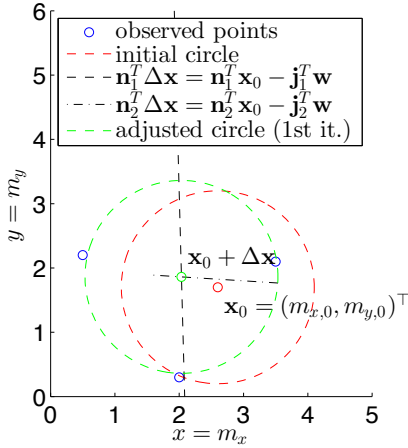


Figure 6: First iteration of the least-squares solution for the circle fit of Figure 4. The updated centre point (green circle) is found as the intersection of the black dotted line and the black dash-dotted line, each representing one line of the normal equation system.

One could think of extending the parabola example of the previous section to a three-dimensional adjustment problem: the third parameter could e.g. be the offset d of an x -parallel directrix l to the x axis; the horizontal projection of the solution $\mathbf{x}_0 + \Delta \mathbf{x}$ would result in the focus \mathbf{f} while the directrix is the parallel projection of the line intersection the planes $x = 0$ and the plane $d = d_0 + \Delta d$ onto the $(f_x, f_y) = (xy)$ plane; the corresponding projection vector is $(0, -\sqrt{2}/2, \sqrt{2}/2)^\top$.

4. HOW TO TREAT MORE THAN THREE DIMENSIONS?

As there is also a descriptive geometry in a Euclidean space of four dimensions [7], the approach given in the previous sections may even be extended to adjustment problems where four parameters x_1, \dots, x_4 are to be determined. Visualization by means of a multi-view orthographic projection may be given as two main ortho-

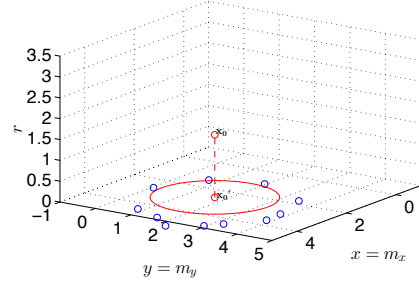


Figure 7: Observed points (blue) and initial estimate \mathbf{x}_0 for the parameters of a circle (red) in the parameter space. The third dimension is the radius r of the circle, so that the length of the red dotted line corresponds to the initial estimate r_0 . The horizontal projection \mathbf{x}'_0 corresponds to the centre of the circle.

nal projections onto the x_1x_2 plane and the x_3x_4 plane, resp., accompanied by two auxiliary orthogonal projections onto the x_2x_3 and the x_4x_1 plane. An example for an adjustment problem in the mentioned configuration may be a sphere where all 4 of its parameters are to be determined. E.g. the coordinates of the center m_x, m_y and m_z may be considered as x_1, x_2 and x_3 , resp. and the observed points can be embedded into this three-dimensional subspace of the parameter space; the fourth dimension x_4 of this parameter space is given by the radius r .

5. STOCHASTICS IN THE CONTEXT OF LEAST-SQUARES ADJUSTMENT

While we have so far treated the observations as constants, each measurement or observation inherits some kind of uncertainty. This is regarded for by assigning a weight matrix P to the observations; the observations are form the vector \mathbf{y} [3, 6]. Therefore, also the observations are updated in each iteration w.r.t. to the least-squares condition

$$\Delta \mathbf{y}^\top P \Delta \mathbf{y} \rightarrow \min.$$

The weight matrix is found as the inverse of the covariance matrix Q_{ll} . Its diagonal elements are

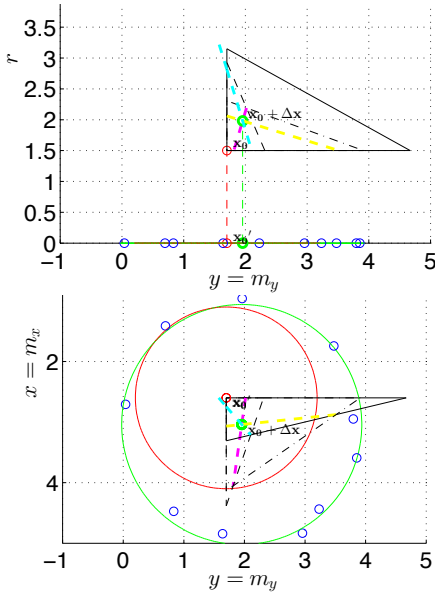


Figure 8: Solution of the normal equation system $\mathbf{x}_0 + \Delta\mathbf{x}$ for the circle parameters retrieved as intersection of the traces of three planes (black triangles). Top: front view, bottom: top view. See Figure 9 for an axonometric visualization. The length of the green dotted line in the front view corresponds to the radius r .

the variances of the corresponding observations while the non-diagonal elements contain the products of the corresponding standard deviations multiplied with the correlation coefficient. Thus, in the case of uncorrelated observations, Q_{II} and P are diagonal matrices.

We retrieve a linearized functional model of the form

$$J_x \Delta\mathbf{x} + J_l \Delta\mathbf{y} + \mathbf{w} = \mathbf{o},$$

also referred to as Gauß-Helmert model [6], and a normal equation system of the kind

$$\begin{pmatrix} J_l Q_{II} J_l^\top & J_x \\ J_x^\top & 0 \end{pmatrix} \begin{pmatrix} \mathbf{k} \\ \Delta\mathbf{x} \end{pmatrix} = \begin{pmatrix} -\mathbf{w} \\ \mathbf{o} \end{pmatrix}. \quad (6)$$

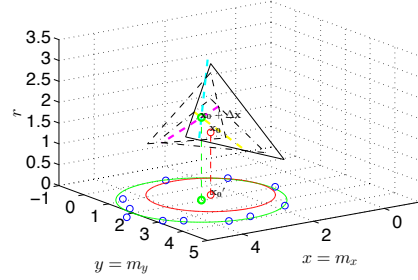


Figure 9: Solution of the normal equation system for the circle parameters in an axonometric view. The horizontal projection of the solution $\mathbf{x}_0 + \Delta\mathbf{x}$ (green circle) gives the centre of the adjusted circle.

The column vector \mathbf{k} contains the Lagrange multipliers, the matrices J_x and J_l are the Jacobian matrices of the condition equations w.r.t. to the unknowns and the observed values, resp.

In case of all observations following a normal distribution, the least-squares adjustment yields the *best linear unbiased estimation* (BLUE) [3, 6].

Using the law of error propagation, the accuracy of the adjusted parameters can also be assessed in a least-squares adjustment. *A-posteriori* analysis of the accuracy of observations is possible via variance component analysis [4]. This analysis additionally allows for determining whether the accuracy estimation was either too optimistic or too pessimistic or adequate.

Uncertainties of the derivatives represented in the Jacobian matrices J_x and J_l may be regarded for in a total least-squares approach aiming for a regularized solution [2].

6. DISCUSSION AND CONCLUSION

The goal of this study was to give a purely geometric motivation and derivation for a least-squares adjustment. In a parameter space of low dimension, dealt with as Euclidean space, the solution can be visualized by means of descriptive geometry if the observations can be embedded into this space. As long as the observations are

treated as constants, the introduction of a stochastic model may be omitted. Examples of two and three parameters were given; the presented approach can in principle be extended to four-dimensional problems.

REFERENCES

- [1] C. F. Gauß. *Theoria Motus Corporum Coelestium in Sectionibus Conicis Solem Ambientium (Theory of the motion of the heavenly bodies moving about the sun in conic sections)*. Cambridge Library Collection – Mathematics. Cambridge University Press, 1809. URL <http://dx.doi.org/10.1017/CBO9780511841705>. Online Publication. Publication Date: April 2012.
- [2] G. H. Golub, P. C. Hansen, and D. P. O’Leary. Tikhonov regularization and total least squares. *SIAM Journal on Matrix Analysis and Applications*, 21(1): 185–194, 1999.
- [3] K.-R. Koch. *Parameter Estimation and Hypothesis Testing in Linear Models*. Springer, 1999.
- [4] K. Kraus. *Photogrammetry – Advanced Methods and Applications*, volume 2. Dümmler Verlag, Bonn, 1997.
- [5] E. M. Mikhail. *Observations And Least Squares*. IEP-A Dun-Donnelley, New York, 1976.
- [6] W. Niemeier. *Ausgleichsrechnung: Eine Einführung für Studierende und Praktiker des Vermessungs- und Geoinformationswesens*. Walter de Gruyter, Berlin – New York, 1st edition, 2002.
- [7] H. Stachel. The right-angle-theorem in four dimensions. *Journal of Theoretical Graphics and Computing*, 3(1): 4–13, 1990.
- [8] W. Wunderlich. *Darstellende Geometrie I*. Hochschultaschenbücher für den Ingenieur.

Bibliographisches Institut, Mannheim, Germany, 1966.

ABOUT THE AUTHOR

Andreas Roncat: born in 1981, studies in geodesy, geometry and computer science at Vienna University of Technology and University of Vienna. Dipl.-Ing. (MSc) 2006. Currently Research Associate and PhD Student at the Research Groups Photogrammetry and Remote Sensing, Department of Geodesy and Geoinformation, Vienna University of Technology. Contact: andreas.roncat@geo.tuwien.ac.at.

THE WORKING DRAWING OF ARCHITECTURE: A TEACHING EXPERIENCE OF REFLECTION ON CODES OF REPRESENTATION

Gianni SAVARRO

School of Architecture and Society, Polytechnic of Milan, Italy

ABSTRACT: The executive representation of the architectural project stimulates questions about the ability of the design to show the materiality of the building. The difficulty of defining the boundary between analogical and symbolic in the use of the graphic sign and the abstractness of the graphic legislation combine to weaken the role of the executive design, dejected either by the excessive realism allowed by the rendering computer, or by the unfavorable comparison with the coding of the technological design.

The contribution exposes the teaching experience gained in the course of the 'representation of the working project', addressed to the students in their final year of the degree course in architecture at the Polytechnic of Milan, which showed the difficulties of the students to overcome the instrumental and decorative conception of the 'design', that is, the difficulty of developing a executive graphic able to reconcile the design choices with the need of transmissibility, without having to give up the 'good design'. The intervention addresses the problematic issue on two levels: the need to define and systematize the teaching of graphic executive codes, adapting them to contemporary standards and technologies, and, on the opposite side, documenting the learning process of the students, whose drawings show the progressive refinement of graphic 'executive' languages, suitable for the formulation of an acceptable synthesis between personal feelings and technological-normative correctness. Academic training in architecture, where the components of composition and design are a priority, the executive drawing is often unable to take on its own identity, standing ambiguously between a large-scale architectural design and a 'mechanical' drawing of the detail, decontextualized however, from the object in its entirety. The themes of the passage of the scale, of the relation analogical/symbolic in the representation of the elements, of the gradual characterization/specialization of the graphic sign with the increase of the scale, identify open issues in the graphical representation, to which the educational experience has tried to provide an answer in terms of systematization of knowledge and building shared with the students of the graphic effective repertoire.

The contribution, by documenting the activities of the laboratory, exposes a teaching trail with lessons of summary classification of graphic executive codes, related to specific technological expertise, led the students' reflection on the need for codification of graphic sign, an indispensable requirement for the transmissibility of the project activity. The recognition of the fragmentation and the lack of coherence of the regulations on the preparation of the final design, often thought to be overcome not only by students, requires a special attention to code, both in the sense of 'set of codes' and the operational dimension of 'coding process'. In contemporary design, more and more dominated by the power of the image, the role of the code as a lexicon of shared reference is essential, both to ensure its update and support the teaching of representation in the task of transmitting a vocabulary, at the same time fixed and evolving.

The contribution proposes to offer, while recognizing the educational limits and the partiality of the results, the material for the reflection of the role of representation recognized

Keywords: Teaching, Executive representation, Graphics Standards, Graphics Education.

1. INTRODUCTION

«An essential element of the design is the formulation of a prescription or model for a finished work, prior to its implementation... There is a real difference between an idea of the design and its implementation» [1] This statement, although derived from the debate focused on the technology of the 60s, centered on the positivistic definition of 'systematic models for designers', confirms its validity in the current context of the working representation of the architecture, in which the technological project can not separate its formalization graphics.

The reference to the formulation of a 'model' refers to the need for a 'code' of the drawing, that is of a representation model that coordinates the design communication with a convenient scheme of reference points. A finished work, because the idea is realized in a consistent forecast, requires a project, and then a drawing: although, sometimes, if you neglect the importance, the *drawing*, that is the set of signs and graphic conventions, is essential to convey not only a 'formal vision of architecture', but to prefigure the materiality of the building.

The graphic code is the element itself of the project, especially in that executive (working project). However you feel, in the teaching of architecture, a certain distance between the importance of the graphic language and its role recognized by educational project, increasingly attracted to contingencies (formal and technological) rather than the essence of the discipline. Of the drawing and not only.

1.1 The problematic context

This reflection on the drawing starts with two elementary considerations, almost banal.

The first refers to the marginal role attributed to the representation, in its various declinations, by the academic training of the architecture: although essential tool of creation and expression, the drawing can with difficulty be separated from the opinion, shared not only among students, of instrumental means, as necessary as side to the project, which underesti-

mates its 'training' dimension of the idea, believing that the information realism is sufficient to its full expression.

The second observation it is the natural consequence. At the marginality of the representation corresponds the superficial understanding of the *drawing* seen as *code* and its deep relationship with the semiotics and the universe of the attribution of meaning to things. The students, deprived of the theoretical wealth of 'drawing', employ so mechanically a language known to the fragmentary character, whose syntax is mistakenly confused with rigid formal codification superseded by the graphic 'modernity'.

The validity of these considerations seems so obvious in the context of the working representation of architecture: the need for a graphical code, essential for sharing and transmitting knowledge, clashes with the approximate knowledge of the instrument itself, supporting the dangerous strengthening of the causes/effects of the decay of the representation.

What solution? The contribution argues, describing a practical learning experience, the need to recover the importance of code to develop a graphics adapted to the new technological and manufacturing requirements, and to computerized means of representation, however able to update the language specification of the drawing with respect the periodic cultural re-configurations.

1.2 Background: the educational project. Limits and potential

The reflection on the working representation of architecture and, more generally, about the meaning, theoretical and operational, of 'graphical code', is gained in the course of 'Representation of the working project', teaching modules, integrated into the 'Laboratory of Construction', aimed at final year students of the Master's program.

The course, activated for two academic years (2011-12, 2012-13) at the School of Architecture and Society of the Polytechnic of Milan, was aimed at defining a working graphical consistent with the technological and

structural deepening required by the laboratory, which offered to students the opportunity to develop, in a comprehensive manner, an architectural project. The laboratory also has not imposed any default line compositional: both for its technological nature and for the synthetic-final function, allowed students to freely express the architectural position acquired throughout the study; this has fostered the emergence of issues related to effectiveness of the drawing (initially minimized by the students), but of interest to the definition of the learning and to initiate a reflection on the significance of the representation and the role of coding to communicate with the (technical) context of building production.

Working in conditions of uncertainty has provided the opportunity to not only make a 'creative' reinterpretation of graphic techniques and the possessed knowledge, but also to create strong interdisciplinary links and greater awareness on codes of representation. Whose new foundation has directed the attention of the course to the nature of the code, to the role of design in the architect's formation, to the history of the representation and, therefore, to the process of giving graphic form to reality.

The partiality of the result reached, however, does not detract from the need to reflect on the drawing: the construction of quality buildings confirms the need to establish appropriate graphic representations; in particular those of the working project, which is the last stage of the design process of the work of architecture, the latest interface between conception and realization.

2. THE CODE OF PROJECT DRAWING. PROBLEMATIC FRAMEWORK

The communication of complex information in a project occurs through multiple means of expression. However, the most effective and historically consolidated means to show the architectural project is the drawing. The word evokes numerous and diverse cultural and graphic images, related to the context in which mature the drawing, to its purpose and the receiver's reception capability. With increasing

specialization of the drawing, the knowledge and the understanding of coded language becomes essential to allow, both to the drawing and the project, adding new elements to the cultural discourse.

The design drawings communicate their information and their larger messages (formal, cultural,...) through the use of different codes, depending on the subsequent levels of processing, the operators of the building process which are addressed ... that express the depth (and complexity) of a specific language itself, that is a set of rules that govern the communication.

It is customary, in fact, to use the term code, borrowed the legal meaning of collection of laws, to confirm the 'prescriptive' and rigid nature of its language. A code, in general, is synonymous with the language, that is a system of signs common to members of a given social body, of a specific language community. The coding is not, of course, only a property of natural languages/out-and-out: all systems of signs, or semiotic systems, are shared codes by a certain community of users with whom you can send various types of messages, combining the signs according to certain rules.

The code of the project drawing is, in this sense, a language.

The graphical code connects meanings and signifiers of a very different nature: «from the projection methods to reproduce the space on a plane, to the geometric simplifications and reductions in scale to the complexity of objects, real or imaginary, to the conventional pattern, the graphics traits and the systems developed for the designation of their coordination» [2]. A real language, in the sense used by De Saussure [3] in his *Cours de linguistique générale*, which probably introduces for the first time the term 'code' in the sense of 'language', characterized by 'words' and 'relationships associations' between the elements (combinations and alternatives) that allow you to structure your thinking (and the drawing too). But what knowledge do students and operators of the building production too, have of this linguistic discipline?

It has the feel of an increasing lack of understanding of the coding among the same professional actors, supported perhaps by the marginalization of the 'drawing' in the academic and professional paths, by the illusory confidence in computing power and the reluctance to the theoretical deepen. The rapid consumption of contemporary culture has not spared the graphic representation, whose deterioration is clearly seen in the working representation. In which, clearly, it shows the need the 'code', but also his limited knowledge. The working project is the decisive moment in which the solution of the problems of the transition from drawing to construction site (from 'theory' to 'practice') is solved by a refined symbolism, apparently contrasting with the operational nature of the drawing.

The update of the 'graphical code' it need not be postponed: the increasing technological change, the regulatory requirements for the award of the construction works, progressively more stringent (for example the 'Code of Public Contracts'), the media requirements of communication of the drawing, require to the codes of the project drawing to upgrade, not only for ensure quality in the project, but to affirm the centrality of the role of graphic representations. The story of 'graphical code' in the present, however, resembles that of the more established codes: as pointed out by the experts of law, codification is often obscured/slowed down by two simultaneous but disparate phenomena: on the one hand, the historical weight of tradition and the other hand the recent trends in the debate on 'decoding vs recoding'.

These phenomena are also found in the graphic field, where the reverence for the history contrasts with the trend, supported by the realism of computer representations of overcoming the conventional linguistic coding. However, even in these guidelines, there are items of note: the decoding identifies, anyway, as the common denominator of the communication process, the need for 'stabilization of the unstable', a consolidation of the language for the effective transmission of messages [4].

The refounding of the code implies, rather than its rewriting, a renewed awareness of its function. And this can be done recovering examples from the history of architecture too. In the words of the authors and critics of architecture, ancient and contemporary, emerges the intertwining among the drawing, the architectural discipline, the practice of design and the construction technology. In the plurality of positions is the close relationship, in all past ages, between the systems of representation and the construction system, such as to determine periodic codifications, that is updates to reflect changing operational needs.

The graphic language, already starting from Leonardo, is emerging as the preferred way to do science and technology: to make a project (in the literal sense of a future projection of an idea) can not be ignored by the codification. And this need is confirmed, further, with the descriptive geometry and with the science of engineering, in which the graphic is not a simple drawing, but the calculation of the structural elements. And then, the graphic code is of decisive importance for the Bauhaus, whose 'art' requires, necessarily, a shared language of expression.

Finally, the need of codification is increased by computer graphics, which radically changing the traditional approach to the drawing of the project, required a major overhaul of the graphic language, especially of the working documents; the invariance of the perceptual window, corresponding to the size of the monitor, the propensity to represent the actual size, greatly reduce, in phase representation, the ultimate control of the project, highlighting problems unknown to the drawing manual [5]. The inability to immediately assess the correctness of the planning and formal contents, heightens the need for codification of design drawing. The late update of the traditional symbolism, inadequate for computer graphics, has led to the phenomenon of abandonment of the code, as well as supported the usefulness of its redefinition, yet another re-establishment of the discipline.

3. THE CODE OF THE PROJECT: REFLECTIONS ON TEACHING EXPERIENCE

The course of 'Representation of the working project' has integrated, for two consecutive academic years, the 'Laboratory of construction', theoretical and application workshop, in the first half of last year's magisterial degree in architecture, which offers students a specific theme of design to be developed from compositional concept to fully definition of the detail construction (fig. 2, on the next page).

3.1 Teaching strategy

The workshop, for its synthetic-final purpose, preferred seminars, collective or in small groups, to the lessons, dedicated to the systematization of (technological and structural) knowledge already acquired.

The exposure of the design documents (drawings and, sometimes, models), the discussion and comparison of the choices made, allowed the group, in its entirety, to assess the progress of the work and to specify, in progress, goals, teaching and planning, to achieve.

The teaching strategy took advantage of the comparison between the different methods of teaching of the three teachers: if, in the first instance, this variety of positions and viewpoints created an apparent disorientation, during the workshop, with the formation of small groups and with the lessons, sometimes addressed to the specific needs of the individual units of work, this 'open formula' has proved to be appropriate to the students and their projects.

3.2 Case study

The identification of a public work, the subject of a competition for students, has been the common denominator of design themes each year. Students were required, although the educational limitations, to respect, in the development of the project, the phases (with relevant documents and drawings) of the legislation on public works (the 'Code of the Public Contracts'), whose implementing regulations establish precisely the project documents.



Figure 1: Cases study: Social housing (above), Town hall (below) (*student work*)

The calendar of activities and seminars of the laboratory has thus traced by the design coded levels: preliminary, definitive and working projects. Although the different cases proposed (Social housing and Town hall, fig. 1), in each year have emerged similar problems and deficiencies, requiring frequent adjustments in the immediate objectives and, more generally, supporting a reflection on the effectiveness of teaching and on the articulation of curricula.

The design project have confirmed the low propensity of architecture students to the technological and structural deepening of the project, as well as the underestimation of the importance of the 'technical' graphic code in the working drawing.

3.3 Student results

Considering the short period of work of an academic semester (little more than four months), the projects of the groups of students have achieved, each with its own way and time,

the intermediate objectives (seminars corresponding to the phases of the preliminary, definitive and working project) and those final, processing of a complete project, both on the technological and structural level, and on the correctness of graphic codes. The groups were called, on the occasion of seminars, to carry out a self-assessment parallel to the evaluation of the teachers. Often the ratings were very similar; sometimes the evaluations differed on about defining the problems, rather than on solutions (coherent) to the wrong analysis.

The student feedback has proved essential for the upgrade planning of the course to the contextual reality: species for the representation, these results have allowed the emergence of the incompleteness of the graphic code, although minimized by students, nearing completion of their studies.

3.4 Code and representation: analytical repertoire of design drawings

The teaching and learning relationship between analog and digital approaches in the process of architectural design has been identified as a priority program of the module of 'representation'. The need to avoid the fruitless coldness of computer design, with its illusory semblance of 'definitive', has helped, in the initial phase of the experience, a manual graphic, progressively implemented by computer representations for the preparation of final materials and of construction detail (fig. 3).

The comparison between the forms of representation has specified the value of the graphic code: it emerged as they are not the scale of reduction (greater or less), the level of precision of the sign (type and thickness of the line) and, in general, all those qualities 'techniques' of the computer drawing, to influence, in a decisive way, the clarity of the drawing. Its readability is given by an appropriate balance between formal correctness and richness of the 'vocabulary', that is the conscious use of a graphical code signifier.

The working representation on the teaching experience: issues, deficiencies and food for thought (fig. 4, 5):



Figure 2: example of a student project: design concept and sequence of drawings

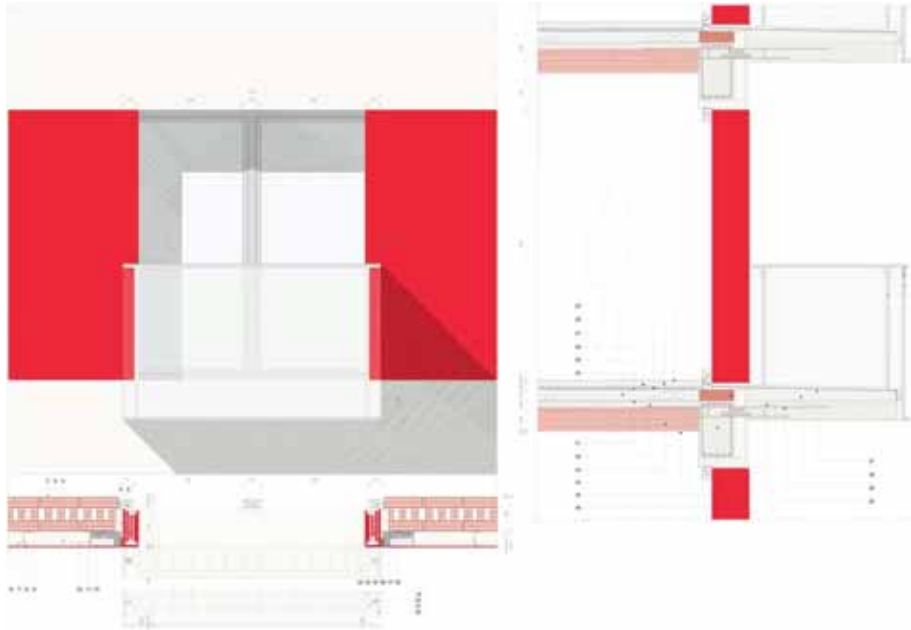


Figure 3: Construction detail: color experiments (*student work*)

- *Ambiguity between graphic code and technical file*: the recurrent inability to adequately represent a technological choice in a whole elaborate, determines the conceptual confusion between the performance specifications of the material and its representation, producing both drawings excessively overloaded with information 'not coded' and, on the contrary, the total absence of drawing, delegating to the technical file communication of the message;
- *-Reinforcing between technological and graphic problems*: the close link between technology and working drawing determines in confusion/clear separation of their goals (satisfaction of quality needs of the building, of the first, coding information, of the second), which causes the mutual reinforcement of communication difficulties and the assignation, by students, of all the problems to only technological aspects, rejecting the idea of a lack of graphic codification.
- *Conception of drawing as a synonym for art*: it's recurrent among architecture students the association of the drawing a form of artistic expression and of visual promotion of the project rather than the consideration of technical and cultural means of transmission of knowledge and information. The graphic representation is seen solely as a means of instrumental driven, in the specific case of the working drawing, by the technology/structural engineering; consequent the lack of interest in deepening the code graph, whose lack of recognition prevents to the same technology to take shape. The erroneous cultural conception of the drawing determines the emptying of its meaning, reduced to exclusive technical means.
- *Analogical vs. symbolic code*: difficulties in

attributing to the signs and forms represented (also due to an outdated manuals to the new building and representation technologies) an analogical or symbolic character. Ambiguity in the use of codes, with the arbitrary invention of signs neither analogical nor symbolic, is not supported by any reference;

- *Analogy as harbor*: the computer hyper-realism and the large-scale encourage the prevalence of an analogical dimension of the drawing, with the aforementioned risk of get next to technical details, decontextualized from the whole;
- *Passage of scale*: the choice of the appropriate scale of reduction, especially in the working drawing, is not obvious; you shows the difficulty of students to use ratios scale appropriate to the understanding of detail, contextualized as a whole, by appropriately varying the graphic techniques.
This has resulted in:
- *Difficulty in selecting the information*: drawings too full or too poor of graphical information, depending on the scale (of print) used, to prevent the immediate reading;
- *Overlap of information*: the revival in all levels of the drawing of the previous information, with the misleading overlap of elements of the architectural/technology /structural design; criticality determined by computerization graphic and by the discontinuance of the separation of meanings;
- *Uncertain thematic of the drawings*: the uncertain selection of the signs, weakened the informative function to the graphic document, devoid of a specific theme. Ex: the design of the heating system associated with the structural scheme rather than to the furniture;
- *Use of mixed messages*: the incoherent use of written information to supplement the lacking graphics, not related to any coding system, is source of contradictions and ambiguities;
- *Graphic uncertainties* in the choice of weights and types line with the change of

scale and of thematic, causes a neutral and inexpressive drawing, though sometimes correct, unable to clearly convey the information at issue.

Recognizing the nature of the graphic difficulties has been the most important achievement of the learning experience, regardless of the actual results. The renovated awareness of the drawing has allowed the partial overcoming of its limits: the observations in progress on the 'drawing' have supported the reflection on the need for the construction of working graphic codes, at the same time personal and conventional, synthesis of 'standards' and graphic 'awareness' individual. To explore the full capacity of the drawing and the information technology in the design process, its primary elements and the role of the combination of graphic signs, has allowed the formulation of a set of ability, a code of the concept, hopefully longer lasting of the educational and planning expedients.

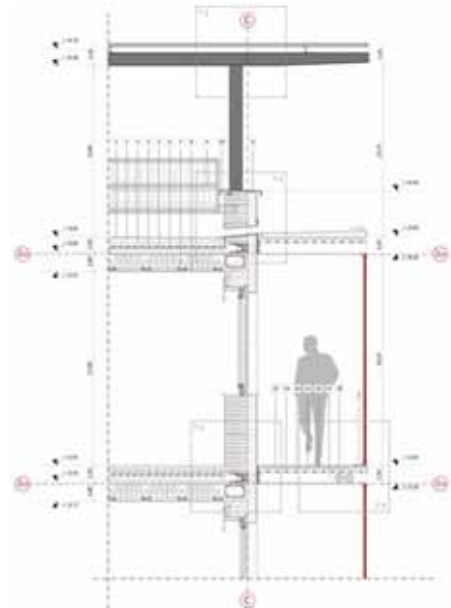


Figure 4: detail of construction: vertical section building

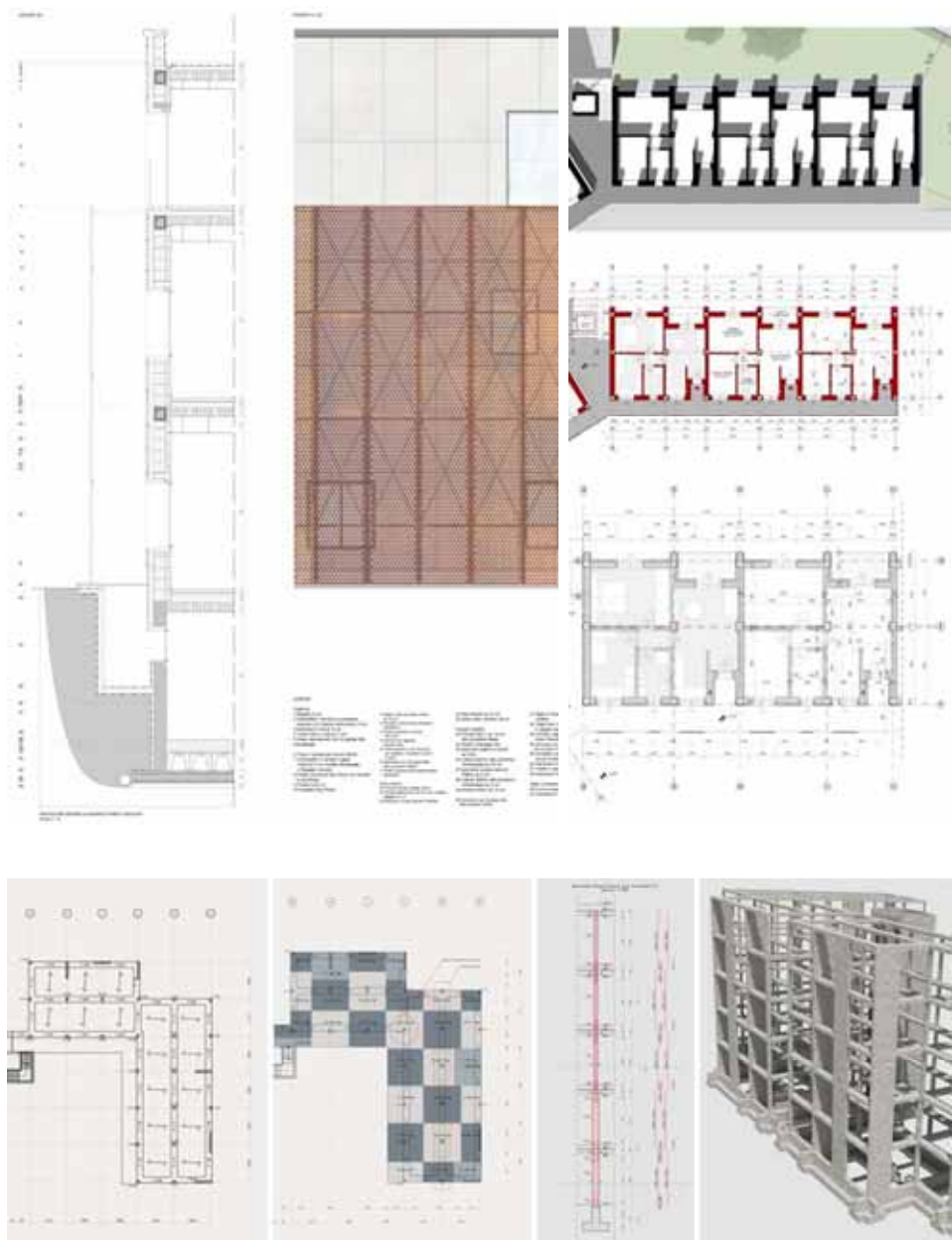


Figure 5: repertoire of drawings laboratory

4. CONCLUSION

"It's sometimes of great help and consolation to preserve of the past a summary representation, marked by a streamlined but reasonable story structure, making it possible to recognize in the present the emergence of a whole new situation" [6]. This is the spirit in which you offer this small and marginal contribution to the debate on 'codes of representation of architecture', aware of the richness of disciplinary tradition (both recent and past) and of the ability of survival of the 'code', although the new complexities.

The teaching experience has demonstrated the impossibility of effective representation without the awareness of the code, confirming the need for its updating, because the analogy can not replace, especially in the contemporary context of building, the symbolism of the codification. It is so emerged the theme of the mutual relationship between the 'geometric structures' of the drawing and the semiotic system of the 'graphic codes', as the graphic model is depending of the semiotic potential of to become image of the phenomenal reality. And vice versa: as in the geometry of sign appear the syntactic structure of the code.

The teaching experience has allowed students and teachers to reflect what the drawing is not only an ancillary means to the project, but by the coded representation is expressed the project and takes a shape. The expression of the idea necessarily requires rules (codes) that are shared to ensure transferability to the discipline (from *discipline*: that is, described by a rule) and not let it fall to a form of artistic whim.

The epistemological nature of the code, its character, fixed and in evolution, perhaps is the only certainty that the drawing possesses for the cultural evolution of the graphic communications.

REFERENCES

- [1] Consideration of Bruno Archer expressed in 1963. Mentioned: L. Pratelli (ed). *Codici del disegno di progetto*. Forum editrice, Udine, 2006, p. 315.

- [2] P. Boltri. *Rappresentazione delle qualità del progetto*. Forum editrice, Udine, 2006, p. 316.
- [3] De Saussure. *Cours de linguistique générale*, Paris, 1922. It. transl. T. De Mauro. *Corso di linguistica generale*. Laterza, Roma-Bari, 1967, p.24.
- [4] See: P. Cappellini. Voce *Codificazione*, in *Enciclopedia Filosofica*. Bompiani, Milano, 2006, vol. 3, p. 1983.
- [5] See: P. Giandebiaggi and G. Ceiner. *Disegno automatico? Riflessione sulle tecniche informatiche di rappresentazione grafica*, in *Presenza tecnica*, Parma, n. 4, luglio-agosto 1997.
- [6] E. Benvenuto. *Architetti, ingegneri e matematici*, in *Casabella*, n.542/543, gennaio-febbraio 1998, p. 16.

ACKNOWLEDGEMENTS

I would like to thank my colleagues in the Laboratory of Construction, prof. Roberto Bolici (course technology of architecture) and Giacomo Boffi (structural engineering), with whom I shared the teaching project.

A special thanks to prof. Luigi Cocchiarella for the discussion on the graphic codes.

All student projects shown here were originated at the Polytechnic of Milan, Italy. Thanks to my hard-working students of the projects shown here: Group A (C. Bonavetti, M. Cittadini, C. D'Orazi, G. Venturini), Group B (M. Filippini, R. Leto, F. Vanoni, S. Wegmann), Group C (C. Franchini, F. Laici, F. Zanetti), Group D (F. Remelli, M. Rossi, N. Sorio) and Group E (A. Fornasari, M. Lecce, S. Marmiroli, C. Torreggiani).

ABOUT THE AUTHOR

Gianni Savarro, Architect and PhD, has been a teacher of technology at secondary school and, since 2011, adjunct professor of drawing at the Polytechnic of Milan.

His research interests are the relationship between the architecture and its representation. He can be reached by e-mail: gsavarr@tin.it or gianni.savarro@polimi.it.

Author Index

- Abdelfattah, Hazem, 1298
Abe, Hirokazu, 355, 363, 600, 1184
Alkhaldi, Ali, 912
Alonso, Alejandro, 103
Alsarhan, Thanaa, 1298
Amoruso, Giuseppe, 936
Ando, Naomi, 1206
Angeles, Jorge, 1192
Arai, Sayaka, 336
- Baglioni, Leonardo, 54
Barczik, Günter, 706
Barej, Martin, 266
Barrallo, Javier, 437
Bauch, Jürgen, 580
Beban-Brkić, Jelena, 690
Bertacchi, Silvia, 130, 548
Blaschitz, Bernhard, 484
Böke, Jens, 1222
Bornemann, Martin, 93
Božić, Ivana, 25
Brailov, Aleksandr Yurievich, 444
Brakhage, Karl-Heinz, 291
Branoff, Theodore, 1174
Buratti, Giorgio, 1236
- Caglioti, Guiseppe, 343
Calisi, Daniele, 245
Cardone, Vito, 847
Ceron, Enrique, 634
Cervantes-Sánchez, J. Jesús, 726
Chiarenza, Stefano, 538
Chizhik, Margarita, 784
Cho, Satoshi, 774
- Choi, Min-Hyung, 103
Ciminieri, Daniele, 321
Cipriani, Luca, 560
Ciuccarelli, Paolo, 321
Čmelková, Viera, 18
Cocchiarella, Luigi, 343, 944
Colagreco, Graciela, 1283
Correia, José V., 405
Corves, Burkhard, 266
Costa, Manuel C. da, 405
Croft Jr., Frank M., 614, 1130, 1142
Cucakovic, Aleksandar A., 262, 1075
Cutellic, Pierre, 10
Czyńska, Klara, 519
- De Spinadel, Vera W., 115, 459
Deng, Li, 285
Denysova, Tetiana, 142
Devahastin Na Ayudhya, Chinnawat,
672
Dobelis, Modris, 1174
Doli, Flamur, 495
Đorđević, Đorđe, 45
Dos Santos, Sérgio Leandro, 643
Dragovic, Magdalena, 1075
Du, Qilu, 1257
Du, Tingna, 1257
Đukanović, Gordana, 45, 187
Dunham, Douglas, 928
Dvoretzky, Alexander, 142
Dzwierzynska, Jolanta, 1031
- Enghardt, Stefan, 580

Fadon, Fernando, 634
Fadon, Laida, 634
Falcón-Morales, Luis E., 865
Fallavollita, Federico, 809
Fan, Qiong, 1043
Fantini, Filippo, 560
Felbrich, Benjamin, 383
Figliolini, Giorgio, 1192
Filipowski, Szymon, 1125
Frick, Ursula, 10
Fritsche, Niels-Christian, 1012
Fu, Yuanguang, 285
Fukui, Yukio, 328

González-Quintial, Francisco, 437
Gorjanc, Sonja, 697, 1168
Grabner, Thomas, 10
Guan, Lijie, 1061
Guerreiro, Ana S., 405
Gumen, Olena, 513
Guo, Feng, 1061

Halas, Helena, 697, 739
Han, Baoling, 1055
Hao, Shilun, 1130
Hast, Anders, 626
He, Yuanjun, 371, 1055
Hemmerling, Marco, 1222
Henschel, Frank, 580
Higashi, Masatake, 893
Hiraoka, Hiroyuki, 607
Hirose, Kenichi, 655, 1048
Hoffmann, Susanne, 266
Hongo, Takeshi, 1226
Honma, Iwao, 310
Hu, Chunxia, 1218
Hünniger, Martin, 197

Iida, Naoki, 655, 1048
Ilić, Maja, 277
Ilyasova, Olga, 784
Imai, Yusuke, 607
Inoue, Aya, 1340
Ishii, Shota, 1206
Itskovich, Elizabeth, 1117
Iwasawa, Shun, 467

Jankowski, Hanna, 1328
Jia, Aichen, 778
Jovic, Biljana S., 262
Jurkin, Ema, 697, 1168

Kager, Helmut, 590
Kanamori, Yoshihiro, 328
Kaneko, Shinya, 840
Kanematsu, Yoshihisa, 167
Kanongchaiyos, Pizzanu, 672
Kasatkina, Elena, 1316
Katić Žlepalo, Mirela, 257
Kato, Michio, 159
Katona, János, 1345
Kaurkin, Viktor, 1316
Kawaguchi, Yoichiro, 467
Kawaharada, Hiroshi, 607
Kaygorodtseva, Natalya, 758
Kaygorodtseva, Tatyana, 758
Kharabayev, Yuriy N., 1291
Kiss-György, Rita, 1159
Klawitter, Daniel, 873
Klett, Yves, 719
Knief, Christian, 266
Koba, Ryotaro, 29
Kobayashi, Masakazu, 893
Kondo, Kunio, 167
Kong, Chuipin, 314

Konishi, Toshimasa, 840
Korchagin, Denis S., 428
Kotarska-Lewandowska, Bożena, 1310
Kovačević, Nikolina, 739
Kuhlmann, Tobias, 712
Kulig, Anna, 572
Kusar, Domen, 1248
Kuwabara, Kazunori, 310
Kuznetsova, Iryna, 1295

Lapaine, Miljenko, 764
Lazarevic, Luka, 1075
Leopold, Cornelië, 956
Li, Gang, 285
Li, Mingqiu, 1218
Li, Shu, 285
Liu, Huilai, 1061
Liu, Wei, 314
Lordick, Daniel, 93, 383, 712
Luo, Xiao, 1055
Lyashkov, Aleksey, 475

Ma, Yan, 285
Makuteniene, Daiva, 820
Maleczek, Rupert, 10
Manevich, Michael, 1117
Mao, Qing-qing, 1257
Marabotto, Maria Paola, 1082
Marchetti, Elena, 732
Maresch, Günter, 301
Maruya, Kazushi, 1320
Masserano, Silvia, 936
Matsumoto, Takuya, 355
Mazzucchi, Alessio, 885
Mele, Giampiero, 1236
Melzer, Sebastian, 93
Ménard, Raphaël, 321

Messina, Barbara, 847
Meunier, Guillaume, 321
Mick, Sybille, 415
Middendorf, Peter, 719
Migliari, Riccardo, 54, 809, 968
Mikami, Koji, 167
Mileikovskiy, Viktor, 513
Milin-Šipuš, Željka, 1087, 1272
Mišić, Slobodan, 45, 187
Mitani, Jun, 328
Mollicone, Antonio, 752
Molnár, Emil, 1345
Mori, Masayuki, 655, 1048
Motegi, Ryuta, 167

Nagakura, Takehiko, 75, 423
Nagashima, Shinobu, 1154
Nanetti, Luca, 982
Nawratil, Georg, 204
Nedeljkovic, Djordje, 1075
Ningrong, Zhang, 1037
Nishii, Misako, 1002
Niteyskiy, Anton S., 216
Niu, Qiang, 314
Noennig, Jörg Rainer, 383
Nosulya, Svetlana, 802

Obradović, Marija Đ., 45, 187
Odehnal, Boris, 235
Ohtani, Tomoko, 1320
Ortale, Horacio, 1283
Otsu, Tatsuo, 336

Palestini, Caterina, 505
Panchuk, Konstantin L., 216, 428
Parque, Victor, 893
Parrinello, Sandro, 548

Pavillet, Axel, 680
Pech, Pavel, 1
Pereda, Raul, 634
Perez, Francisco, 794
Pfurner, Martin, 1024
Pisacane, Nicola, 661
Pletenac, Lidija, 1272
Primorac Gajčić, Ljiljana, 1087
Prok, István, 1345

Qinghua, Xie, 1037

Rachkovskaya, Galina S., 1291
Rachkovskaya, Natalya S., 1291
Renken, Remco J., 982
Rickenbacher, Martin, 590
Romão, Luís, 405
Romor, Jessica, 968
Roncat, Andreas, 590, 1355
Rosales, Enrique A., 865
Röschel, Otto, 903
Rossi-Costa, Luisa, 732
Rovo, Federica, 1236
Ruan, Pingping, 778
Rubinowicz, Pawel, 393
Rutten, David, 826

Saito, Takafumi, 1002
Saiz, Francisco, 634
Sakaki, Ai, 29
Salvatore, Marta, 809
Sánchez-Parandiet, Antonio, 437
Sato, Hisashi, 774
Savarro, Gianni, 1362
Schadlbauer, Josef, 1024
Schröcker, Hans-Peter, 1024
Sdegno, Alberto, 936

Sejfried, Michael, 749
Shelomovskii, Vladimir, 802, 1095
Shibata, Akihiro, 1206
Shiina, Kumiko, 336
Shinya, Mikio, 1340
Shiraishi, Michio, 1340
Shonoda, Emad, 121
Shvalb, Nir, 1117
Šimić Horvath, Marija, 690
Sinha, Shriprakash, 982
Sliepčević, Ana, 25, 739
Sokas, Algirdas, 531
Sripian, Peeraya, 81
Stachel, Hellmuth, 1192
Stanberry, Larissa, 1328
Stavrić, Milena, 277, 1066
Stepanov, Yuriy, 1316
Suzuki, Hirotaka, 224
Suzuki, Kenjiro, 1226, 1264
Szirmai, Jenő, 1345

Tachi, Tomohiro, 66
Takahashi, Akira, 363
Takahashi, Kuniko, 66
Tan, Adrian H., 614, 1130, 1142
Tan, Fabian H., 614, 1130, 1142
Tanaka, Ichiro, 1111
Taras, Iryna, 922
Teixeira, Fábio Gonçalves, 643
Teofilovic, Natasa K., 262
ter Horst, Gert J., 982
Tominaga, Hiroki, 655, 1048
Tong, Bingshu, 1055
Trautz, Martin, 266
Tsuchida, Takahiro, 167
Tsuruta, Naoya, 328
Tsutsumi, Emiko, 1226

Valencia García, Germán, 858

Velichová, Daniela, 149

Vidak, Stipe, 375

Volgemut, Mateja, 1248

Volkov, Vladimir, 475, 784

Wada, Kazuma, 1184

Wang, Chunhua, 1061

Wang, Lang, 1043

Wang, Niannian, 778

Wang, Yongxing, 371

Wang, Ziru, 1218

Weiß, Gunter, 115, 121, 415, 1283

Wiesenhütter, Sebastian, 383

Wildberger, Norman, 912

Wiltsche, Albert, 1066

Yamade, Miya, 600

Yamaguchi, Yasushi, 66, 81

Yamahata, Nobuhiro, 1206

Yamamoto, Toshikazu, 1226

Yang, Jie, 1055

Yang, Wei, 1043

Yano, Hiroshi, 1226

Yasufuku, Kensuke, 38, 355

Yoshii, Hideki, 774

Yu, Haiyuan, 371

Zawidzki, Machi, 75, 423

Zhang, Jing-Ying, 1043

Zhou, Xionghui, 314

Zsombor-Murray, Paul, 726

Zwoliński, Adam, 175