











    **SIGGRAPH**   
SAN ANTONIO **Conference Abstracts and Applications**

Computer Graphics Annual Conference Series, 2002          
A Publication of ACM SIGGRAPH



# Conference Abstracts and Applications

**COMMUNITY OUTREACH** | **Nancy Wood** | San Antonio College

**EDUCATORS PROGRAM** | **James L. Mohler** | Purdue University

**EMERGING TECHNOLOGIES** | **Scott Senften** | SGI

**PANELS** | **Christopher D. Shaw** | Georgia Institute of Technology

**SIGKIDS** | **Marc J. Barr** | Middle Tennessee State University

**SKETCHES & APPLICATIONS** | **Doug Roble** | Digital Domain

**SPECIAL SESSIONS** | **Jackie White** | California State University - Los Angeles

**STUDIO** | **Dan Collins** | Arizona State University / PRISM

**WEB GRAPHICS** | **Simon Allardice** | clingfish.com



SIGGRAPH 2002 Conference Abstracts and Applications

Computer Graphics Annual Conference Series, 2002

The Association for Computing Machinery, Inc.  
1515 Broadway  
New York, New York 10036 USA

ISBN 1-58113-524-6  
ISSN 1098-6138  
ACM Order No. 435023

Additional copies may be ordered pre-paid from:  
ACM Order Department  
P.O. Box 12114  
Church Street Station  
New York, New York 10257 USA

Or, for information on accepted European currencies  
and exchange rates, contact:  
ACM European Service Center  
108 Cowley Road  
Oxford OX4 1JF  
United Kingdom  
+44.1.865.382338  
+44.1.865.381338 fax

Credit card orders from U.S. and Canada:  
800.342.6626

Credit card orders from the New York  
metropolitan area and outside the U.S.:  
+1.212.626.0500

Single copy orders placed by fax:  
+1.212.944.1318

Credit card orders may also be placed by mail.

Electronic mail inquiries may be directed to:  
[orders@acm.org](mailto:orders@acm.org)

Please include your ACM member number and  
the ACM order number with your order.

**Community Outreach**

Chair: Nancy Wood (San Antonio College)

Introduction .....	11
Committee .....	12

**Educators Program**

Chair: James L. Mohler (Purdue University)

Introduction .....	13
Committee .....	14

**Educators Program Forums**

A Knowledge Base for the Computer Graphics Discipline .....	15
<i>Gary R. Bertoline, Cary Laxer</i>	
Game Development, Design & Analysis Curriculum .....	16
<i>Jason Della Rocca, Warren Spector, Eric Zimmerman, Robin Hunike</i>	
Inspiring the Renaissance Person .....	17
<i>Donna Cox</i>	
Teaching Gems for Art and Design .....	18
<i>Dena Elisabeth Eber, Bonnie Mitchell, Heather Elliott</i>	
The Role of Creativity in Computer Graphics Education .....	19
<i>Bruce Wands</i>	
Teaching Gems for Computer Science and Engineering .....	20
<i>Mike Bailey, Steve Cunningham, Lew Hitchner</i>	

**Educators Program Papers**

Problems Using Components in Educational Software .....	21
<i>Anne Morgan Spalter</i>	
What Do Computers Eat? Teaching Beginners to Think Critically About Technology and Art .....	26
<i>Tiffany Holmes</i>	
SIGGRAPH as Textbook: Learning Skills for Undergraduates .....	29
<i>Kevin L. Novins</i>	
Mathematics and Geometry Education With Collaborative Augmented Reality .....	33
<i>Hannes Kaufmann, Dieter Schmalstieg</i>	
Assignment: Scene Graphs in Computer Graphics Courses .....	38
<i>Dennis J. Bouvier</i>	
Building Computer Graphics Education in Developing African Countries ..	42
<i>Sampson D. Asare, Petros M. Mashwama, Steve Cunningham</i>	
Macromedia Flash in Physics Education: ASPIRE's Interactive Online Labs and Lessons .....	45
<i>Julie Callahan, Charles C. H. Jui</i>	
The VERTEX Project: Exploring the Creative Use of Shared 3D Virtual Worlds in the Primary (K-12) Classroom .....	48
<i>Fiona Bailey, Magnus Moar</i>	

**Educators Program**

Chair: James L. Mohler (Purdue University)

**Educators Program Panels**

K-12 and Industry Partnering .....	51
<i>Darlene Wolfe, Timothy Comolli, Jeff Scheetz, Roger Cotton, Chris Stapleton</i>	
Animating Art History for Teaching .....	53
<i>Roberta Tarbell, LiQin Tan</i>	

**Educators Program Workshops**

Building an Affordable Projective, Immersive Display .....	55
<i>Dave Pape, Josephine Anstey</i>	
Life Drawing and 3D Figure Modeling With Maya .....	56
<i>Gregory P. Garvey</i>	
Integrating Web 3D Into 3D Animation Curricula .....	57
<i>Mitch Williams</i>	
Hi Tech-Lo Tech: K-12 Science Visualization .....	58
<i>Susana Maria Halpine</i>	
Teaching Human Facial Modeling Through Plaster Face Casting .....	59
<i>Adam Watkins</i>	

**Emerging Technologies**

Chair: Scott Senften (SGI)

Introduction .....	60
Committee .....	61
Tomorrow's Yesterday: Mapping the E-Tech Continuum .....	62
<i>John M. Fujii</i>	
A New Step-in-Place Locomotion Interface for Virtual Environment With Large Display System .....	63
<i>Laroussi Bouguila, Makoto Sato, Shoichi Hasegawa, Hashimoto Naoki, Naoki Matsumoto, Atsushi Toyama, Jelel Ezzine, Dalel Maghrebi</i>	
ARS BOX with Palmist - Advanced VR-System based on commodity hardware .....	64
<i>Horst Hörtnner, Christopher Lindinger, Robert Praxmarer, Andreas Riedler</i>	
Augmenting the Reality with 3D Sound Sources .....	65
<i>Michael Haller, Daniel Dobler, Philipp Stampfl</i>	
Audio Haptics .....	66
<i>Hiroo Iwata, Hiroaki Yano, Hiromi Igawa</i>	
Block Jam .....	67
<i>Henry Newton-Dunn, Hiroaki Nakano, James Gibson</i>	
Cyberarium Knowledge Fountain .....	68
<i>David Warner, Matt Carbone</i>	
Distributed Systems of Self-reconfiguring Robots .....	69
<i>Zach Butler, Robert Fitch, Keith Kotay, Daniela Rus</i>	
Focus Plus Context Screens: Visual Context and Immersion on the Desktop	70
<i>Patrick Baudisch, Nathan Good</i>	

## Emerging Technologies

Chair: Scott Senften (SGI)

Immersive and Interactive Rear-Projected Stereo DLP BR Center . . . . .	71
<i>Andrew Joel, Michael Brown, Almos Elekes, Paul Mlyniec, Malte Zoeckler</i>	
Lewis the Robotic Photographer . . . . .	72
<i>Cindy M. Grimm, William D. Smart</i>	
NONA-Vision . . . . .	73
<i>Hiroo Iwata, Hiroaki Yano, Motohiro Tsuzuki, Fumitaka Nakaizumi, Takayuki Yoshioka, Yutaka Miyakita</i>	
Occlusive Optical See-through Displays in a Collaborative Setup . . . . .	74
<i>Kiyoshi Kiyokawa, Hiroyuki Ohno, Yoshinori Kurata</i>	
Physiological Reaction and Presence in Stressful Virtual Environments . . . . .	75
<i>Michael Meehan, Mary Whitton, Sharif Razzaque, Paul Zimmons, Brent Insko, Greg Combe, Ben Lok, Thorsten Scheuermann, Samir Naik, Jason Jerald, Mark Harris, Angus Antley, Frederick P. Brooks, Jr.</i>	
Public Anemone: An Organic Robot Creature . . . . .	76
<i>Cynthia Breazeal, Andrew Brooks, Matt Hancher, Cory Kidd, John McBean, Dan Stiehl, Josh Strickon</i>	
Regeneration of Real Objects in the Real World . . . . .	77
<i>Hiroto Matsuoka, Akira Onozawa, Hidenori Sato, Hisao Nojima</i>	
SmartFinger : Nail-Mounted Tactile Display . . . . .	78
<i>Hideyuki Ando, Takeshi Miki, Masahiko Inami, Taro Maeda</i>	
The Interactive Window . . . . .	79
<i>Joseph A. Paradiso, Che King Leo, Nicholas Yu, Marc Downie</i>	
The Virtual Showcase: A Projection-Based Multi-User Augmented Reality Display . . . . .	80
<i>Oliver Bimber, Bernd Fröhlich, Dieter Schmalstieg, L. Miguel Encarnação</i>	
TWISTER: A Media Booth . . . . .	81
<i>Kenji Tanaka, Junya Hayashi, Yutaka Kunita, Masahiko Inami, Taro Maeda, Susumu Tachi</i>	
Ultrasound Visualization with the Sonic Flashlight . . . . .	82
<i>Damion Shelton, George Stetten, Wilson Chang</i>	
Virtual Chanbara . . . . .	83
<i>Daijiro Koga, Takahiro Itagaki</i>	

## Panels

Chair: Christopher Shaw (Georgia Institute of Technology)

Introduction . . . . .	84
Committee . . . . .	85
When Will Ray-Tracing Replace Rasterization? . . . . .	86
<i>Brad Grantham, Kurt Akeley, David Kirk, Larry Seiler, Philipp Slusallek</i>	
Digital Humans: What Roles Will They Play? . . . . .	88
<i>Norman I. Badler, Phil LoPiccolo, Nadia Magnenat-Thalmann, Laurie McColloch, Evan Marc Hirsch, Athomas Goldberg</i>	
Extending Interface Practice: An Ecosystem Approach . . . . .	90
<i>Andruid Kerne, Natalie Jeremijenko, Thecla Schiphorst, Michael Mateas, Wolfgang Strauss, Will Wright</i>	

## Panels

Chair: Christopher Shaw (Georgia Institute of Technology)

The Future of Computer Graphics: An Enabling Technology? . . . . .	93
<i>David S. Ebert, Bill Buxton, Patricia Davies, Elliot K. Fishman, Andrew Glassner</i>	
Symposium on Computer Animation in Fast Forward . . . . .	95
<i>Michael Gleicher, Michael Cohen, Nancy Pollard, Jessica Hodgins, Michiel van de Panne</i>	
The Demo Scene . . . . .	96
<i>Vincent Scheib, Theo Engell-Nielsen, Saku Lehtinen, Eric Haines, Phil Taylor</i>	
Animation's Turning Tide . . . . .	98
<i>Matt Elson, Eamonn Butler, Scott Clark, Carlos Saldanha</i>	
Unsolved Problems in Mobile Computer Graphics and Interaction . . . . .	99
<i>Mark Ollila, Staffan Björk, Kevin Bradshaw, Steven Feiner, Kari Pulli</i>	
Graphics in the Large: Is Bigger Better? . . . . .	101
<i>David J. Kasik, Loren Carpenter, Brian Fisher, Richard A. May, Norbert Streitz</i>	
How Does Motion Capture Affect Animation? . . . . .	103
<i>Suba Varadarajan, Barb Helfer, Margaret S. Geroch, Evan Hirsch, Joan Staveley, Tom Tolles</i>	
Interactive Stories: Real Systems, Three Solutions . . . . .	105
<i>Noah Wardrip-Fruin, Andrew Stern, Peter Molyneux, Michael Mateas, Bernard Yee</i>	
Games: the Dominant Medium of the Future . . . . .	107
<i>Bob Nicolli, Ken Perlin, Glenn Entis, Patrick Gilmore, J. C. Herz, Alex Pham, Will Wright</i>	
<b>sigKIDS</b>	
Chair: Marc Barr (Middle Tennessee State University)	
Introduction . . . . .	109
Committee . . . . .	110
Interactive Animation as an Educational Tool in "Winter Dreams" . . . . .	111
<i>Daria Tsoupikova</i>	
Gaming as an Educational Tool: Internet Scavenger Hunt . . . . .	113
<i>Karen Monahan</i>	
FORM . . . . .	114
<i>Hilary J. Wright</i>	
The Floating Words for Kids . . . . .	115
<i>Satoko Moroi, Shinji Sasada, Ryoji Shibata</i>	
Toys to Teach - Mathematics as a Collaborative Climbing Exercise . . . . .	117
<i>James Dai, Michael Wu, Jonathan Cohen, Troy Wu, Maria Klawe</i>	
Jollie-Mail Playground . . . . .	120
<i>Yuichiro Kamata</i>	
The Virtual Dig . . . . .	122
<i>Robert Dunn</i>	
Virtual Studio: Virtual Reality in Art . . . . .	124
<i>DaShawn L. Hall, Kenneth Sakatani</i>	

## sigKIDS

Chair: Marc Barr (Middle Tennessee State University)

An Application of Tangible Interfaces in Collaborative Learning Environments .....125  
*Lori Scarlatos*

GollyGee Blocks™: A 3D Modeler for Children .....127  
*Jonathan T. Blocksom*

Anansi's World of Folklore .....128  
*Jacqueline Nuwame*

The ToyScout's Immersive Jukebox .....129  
*Christopher Stapleton*

## Sketches & Applications

Chair: Doug Roble (Digital Domain)

Introduction .....131

Committee .....132

2D/3D Hybrid Character Animation on "Spirit" .....133  
*Doug Cooper*

3D Browser for Interactive Television .....134  
*Kotro Petri, Mannerkoski Olli, Lesonen Hannu, Lustila Risto*

3D Haptic Shape Perception Using a 2D Device .....135  
*Huirong Han, Juli Yamashita, Issei Fujishiro*

3D Layout and Propagation of Environmental Phenomena for *Ice Age* ...136  
*Kevin Thomason, Robert V. Cavalieri*

3D Reconstruction of Walking Behaviour Using a Single Camera .....137  
*Arunachalam Somasundaram, Rick Parent*

A 2D Sketch Interface for a 3D Model Search Engine .....138  
*Patrick Min, Joyce Chen, Thomas Funkhouser*

A Dynamic Motion Control Middleware for Computer Games .....139  
*Masaki Oshita, Akifumi Makinouchi*

A Handheld Virtual Mirror .....140  
*Alexandre François, Elaine Kang, Umberto Malesci*

A new Interface for the Virtual World Foot Motion Sensing Input Device ..141  
*Barrera Salvador, Hiroki Takahashi, Masayuki Nakajima*

A Phenomenological Model for Bokeh Rendering .....142  
*Juan Buhler, Dan Wexler*

A Semiotic Approach to Narrative Manipulation .....143  
*Maria Alberta Alberti, Dario Maggiorini, Paola Trapani*

A Virtual Reconstruction of the Cone Sisters' Apartments .....144  
*Alan Price, Dan Bailey*

Accurate Image Based Re-lighting through Optimization .....145  
*Pieter Peers, Philip Dutré*

An Inexpensive 3D Camera .....146  
*Askold Strat, Manuel M. Oliveira*

Art and Education using Direct Manipulation of a Sensor Array .....147  
*Taly Sharon*

## Sketches & Applications

Chair: Doug Roble (Digital Domain)

ASR - Augmented Sound Reality .....148  
*Daniel Dobler, Michael Haller, Philipp Stampfl*

Automatic Generation of Pencil Drawing Using LIC .....149  
*Xiaoyang Mao, Yoshinori Nagasaka, Atsumi Imamiya*

Backseat Gaming: Augmented Reality with Speed .....150  
*Liselott Brunnberg, Mark Ollila*

BioMorphic Typography .....151  
*Diane Gromala*

Body Building Through Weight Training: Using Fitting Techniques for Skin Animation .....152  
*Xiaohuan Corina Wang, Cary Phillips*

Bringing Computer Graphics to Everyday Environments with Informative Art .....153  
*Tobias Skog, Sara Ljungblad, Lars Erik Holmquist*

Calming Visual Spaces: Learning from Kyoto Zen Gardens .....154  
*Michael J. Lyons, Gert J. Van Tonder, Ian Shortreed, Nobuji Tetsutani*

Cameras and Point-of-View in the Gamespace .....155  
*Jay Riddle*

Challenges of the Homeland Pan in "Spirit" .....156  
*Doug Cooper*

Color Transformation Based on the Basic Color Categories of a Painting .157  
*Youngha Chang, Sugura Saito, Masayuki Nakajima*

Computer Generated Clay Animation .....158  
*Daisuke Goto, Junichi Hoshino*

Computer Graphics to Illustrate the Development of a Human Embryo for Professional Medical Education .....159  
*Koh Kakusho, Yutaka Minekura, Michihiko Minoh, Shinobu Mizuta, Tomoko Nakatsu, Kohei Shiota*

Creating 3D Painterly Environments for Disney's "Treasure Planet" .....160  
*Kyle Odermatt, Chris Springfield*

Creation of a Photo-Real CG Human .....161  
*Brian Goldberg*

CT (City Tomography) .....162  
*Fumio Matsumoto, Akira Wakita*

Curvature Driven Sampling of Displacement Maps .....163  
*Johannes Hirche, Alexander Ehlert*

Custom Designs for Digital Imaging on Textiles .....164  
*Bonny Lhotka*

Depth-Complexity Based Occluder Selection .....165  
*Gerhard Kurka*

Detail Calibration for Out-of-Core Model Simplification through Interlaced Sampling .....166  
*Guangzheng Fei, Nadia Magnenat-Thalmann, Kangying Cai, Enhua Wu*

Digital Pyro for *Reign of Fire* .....167  
*Patrick Dalton, Rob Rosenblum, Shyh-Chyuan Huang, Lawrence Lee, Hank Driskill*



## Sketches & Applications

Chair: Doug Roble (Digital Domain)

Diminishing Head-Mounted Display for Shared Augmented Reality . . . . .	168
<i>Masayuki Takemura, Yuichi Ohta</i>	
Diorama Engine - A 3D Video Storyboard Editor for 3D Computer Animation . . . . .	169
<i>Koji Mikami, Toru Tokuhara, Mitsuru Kaneko</i>	
DIVIPRO: Distributed Interactive Virtual PROtotyping . . . . .	170
<i>Mashhuda Glencross, James Marsh, Jon Cook, Sylvain Daubrenet, Steve Pettifer, Roger Hubbard</i>	
DocuDrama Conversations . . . . .	171
<i>Leonie Schäfer, Elaine M. Raybourn, Amanda Oldroyd</i>	
Dragon Scales: The Evolution of ScaleTool for "Reign of Fire" . . . . .	172
<i>Ernest J. Petti, Thomas V Thompson II, Adolph Lusinsky, Hank Driskill</i>	
Drawing With Feeling: Designing Tactile Display for Pen . . . . .	173
<i>Ivan Poupyrev, Shigeaki Maruyama</i>	
Dynamic Simulation of Wing Motion on "Reign of Fire" . . . . .	174
<i>Carlos Gonzalez-Ochoa, David Eberle, Rob Dressel</i>	
Dynamic Skin Deformation and Animation Controls using Maya Cloth for Facial Animation . . . . .	175
<i>Jimmy Chim, Hyunsuk Kim</i>	
Dynamics and Dodos: Rigging and Animation Methods for <i>Ice Age</i> . . . . .	176
<i>Adam Burr, Ross Scroble</i>	
Editable Dynamic Textures . . . . .	177
<i>Gianfranco Doretto, Stefano Soatto</i>	
Embodied Interaction . . . . .	178
<i>Bill Hill</i>	
Evolution of a VFX voxel tool . . . . .	179
<i>Alan Kapler</i>	
Example-Based Interpolation of Human Motion . . . . .	180
<i>Hiroshi Mori, Jun'ichi Hoshino</i>	
Exploiting Screen Space . . . . .	181
<i>Ian Mackinnon</i>	
Expressive Features for Movement Exaggeration . . . . .	182
<i>James W. Davis, Vignesh S. Kannappan</i>	
Firefighter Training Virtual Environment . . . . .	183
<i>Tazama St. Julien, Chris D. Shaw</i>	
Folded: Negotiating The Space Between Real and Virtual Worlds . . . . .	184
<i>Samantha Krukowski</i>	
FlowField: Investigating the Semantics of Caress . . . . .	185
<i>Timothy Chen, Sidney Fels, Thecla Schiphorst</i>	
Foamy Creatures: Digital Domain wrangles whitewater for "Lord of the Rings" . . . . .	186
<i>Markus Kurtz, Greg Duda</i>	
Fracture generation on polygonal meshes using Voronoi polygons . . . . .	187
<i>Saty Raghavachary</i>	

## Sketches & Applications

Chair: Doug Roble (Digital Domain)

Generating Feather Coats using Bézier Curves . . . . .	188
<i>Lisa Streit, Wolfgang Heidrich</i>	
"GEO-COSMOS": The World's First Spherical Display . . . . .	189
<i>Tamotsu Machida, Tsuyoshi "Go" Hotta</i>	
Hardware-accelerated Texture and Edge Antialiasing using FIR Filters . . . . .	190
<i>Koen Meinds, Bart Barenbrug, Frans Peters</i>	
Harmonic 3D Shape Matching . . . . .	191
<i>Michael Kazhdan, Thomas Funkhouser</i>	
Hiding Spaces: a CAVE of Elusive Immateriality . . . . .	192
<i>Cynthia Beth Rubin, Daniel F. Keefe</i>	
High-speed Conversion of Floating Point Images to 8-bit . . . . .	193
<i>Bill Spitzak</i>	
Hover: Conveying Remote Presence . . . . .	194
<i>Dan Maynes-Aminzade, Beng-Kiang Tan, Ken Goulding, Catherine Vaucelle</i>	
How a CSG-Based Raytracer Saves Time: Lighting and Scripting for <i>Ice Age</i> . . . . .	195
<i>David Esneault, Mitch Kopelman, Jodi Whitsel</i>	
How to Dress Like a Jedi: techniques for digital clothing . . . . .	196
<i>Ari Rapkin</i>	
ICARUS: Interactive Reconstruction from Uncalibrated Image Sequences . . . . .	197
<i>Simon Gibson, Jon Cook, Toby Howard, Roger Hubbard</i>	
Image-based Environment Matting . . . . .	198
<i>Yonatan Wexler, Andrew W. Fitzgibbon, Andrew Zisserman</i>	
Image-Based Illumination for Electronic Display of Artistic Paintings . . . . .	199
<i>Da young Ju, Jin-Ho Yoo, Kyoung Chin Seo, Gregory Sharp, Sang Wook Lee</i>	
Implementing the Continuous Staircase Illusion in OpenGL . . . . .	200
<i>Mark W. Scott</i>	
Improving frameless rendering by focusing on change . . . . .	201
<i>Abhinav Dayal, Benjamin Watson, David Luebke</i>	
Integrating Lenticular into Digital Printmaking . . . . .	202
<i>Dorothy Simpson Krause</i>	
Interactive Level-Set Tools for Photo Editing . . . . .	203
<i>Thomas Howard, Bryan Morse</i>	
Interactive Visualization of Large Scale Time Varying Data Sets . . . . .	204
<i>Patric Ljung, Mark Eric Dieckmann, Anders Ynnerman</i>	
Interruptible Rendering . . . . .	205
<i>J. Cliff Woolley, David Luebke, Ben Watson</i>	
Inverse Texture Warping . . . . .	206
<i>Robert Falco, Hank Driskill</i>	
Investigating Face Space . . . . .	207
<i>Steve DiPaola</i>	
Lewis the Robotic Photographer . . . . .	208
<i>Cindy Grimm, William D. Smart</i>	

## Sketches & Applications

Chair: Doug Roble (Digital Domain)

"Loops" - a digital portrait . . . . .	209
<i>Marc Downie</i>	
<i>Lord of the Rings: Animation That Was Not There</i> . . . . .	210
<i>Piotr Karwas</i>	
"Low Level" Intelligence for "Low Level" Animation . . . . .	211
<i>Damian Isla, Bruce Blumberg</i>	
Magical Face: Integrated Tool for Muscle Based Facial Animation . . . . .	212
<i>Tatsuo Yotsukura, Mitsunori Takahashi, Shigeo Morishima, Kazunori Nakamura, Hirokazu Kudoh</i>	
MARE : Multiuser Augmented Reality Environment on table setup . . . . .	213
<i>Raphael Grasset, Jean-Dominique Gascuel</i>	
MasterMotion: Full Body Wireless Virtual Reality for Tai Chi . . . . .	214
<i>Philo Tan Chua, Rebecca Crivella, Bo Daly, Ning Hu, Russ Schaaf, David Ventura, Todd Camill, Jessica Hodgins, Randy Pausch</i>	
Maximum Entropy Light Source Placement . . . . .	215
<i>Stefan Gumhold</i>	
MOCAP Game Reserve: A Study of Puppetry and Motion Capture . . . . .	216
<i>Charlotte Belland</i>	
Modal Analysis for Real-Time Viscoelastic Deformation . . . . .	217
<i>Chen Shen, Kris K. Hauser, Christine M. Gatchalian, James O'Brien</i>	
Modeling the Accumulation of Wind-Driven Snow . . . . .	218
<i>Bryan E. Feldman, James F. O'Brien</i>	
Modified Marching Octahedra for Optimal Regular Meshes . . . . .	219
<i>Hamish Carr, Thomas Theußl, Torsten Möller</i>	
Motion capture done dirt cheap . . . . .	220
<i>Stefan Gustavson</i>	
Motion-Based Shape Illustration . . . . .	221
<i>Eric B. Lum, Aleksander Stoppel, Kwan-Liu Ma</i>	
Music-Driven Motion Editing . . . . .	222
<i>Marc Cardle, Stephen Brooks, Loic Barthe, Mo Hassan, Peter Robinson</i>	
Now and then, here and there - Industrilandskapet . . . . .	223
<i>Konstantin Economou, Mark Ollila, Martin Etherton Friberg, Anders Ynnerman</i>	
Occidio . . . . .	224
<i>Timothy Nohe</i>	
Painterly Fire . . . . .	225
<i>Saty Raghavachary, Fernando Benitez</i>	
Painting with Light . . . . .	226
<i>Kevin G. Suffern</i>	
Pastel-like Rendering Considering the Properties of Pigments and Support Medium . . . . .	227
<i>Kyoko Murakami, Reiji Tsurono</i>	
Perceptual Gaze Extent & Level Of Detail in VR: Looking Outside the Box .228	
<i>Hunter Murphy, Andrew T. Duchowski</i>	

## Sketches & Applications

Chair: Doug Roble (Digital Domain)

Perceptual Tone Mapping Operators for High Dynamic Range Scenes . . .229	
<i>Patrick Ledda, Alan Chalmers, Greg Ward</i>	
Performance driven computer graphics making <i>Odyssey</i> . . . . .	230
<i>Markus Manninen</i>	
Pop.ey.e: A Pop-out Video Camera System for Personal Use . . . . .	231
<i>Isao Mihara, Takahiro Harashima, Shunichi Numazaki, Miwako Doi</i>	
Probabilistically Placing Primitives . . . . .	232
<i>Adrian Secord, Wolfgang Heidrich</i>	
Probability Paint: Controlling Group Characteristics with PDFs . . . . .	233
<i>Kathleen Gretchen Greene</i>	
Pushing the limits of L-systems for time-lapse vine growth in "The Time Machine" . . . . .	234
<i>Jonah Hall</i>	
Quikwriting on the Responsive Workbench . . . . .	235
<i>Jerome Grosjean, Sabine Coquillart</i>	
RAPTOR: Towards Augmented Paleontology . . . . .	236
<i>Oliver Bimber, L. Miguel Encarnação</i>	
Real Time Video Effects on a PlayStation2 . . . . .	237
<i>Sarah Witt</i>	
Real-time 3D interaction for Augmented and Virtual Reality . . . . .	238
<i>Simon Prince, Adrian David Cheok, Farzam Farbiz, Todd Williamson, Nik Johnson, Mark Billingham, Hirokazu Kato</i>	
Real-Time Image-Space Outlining for Non-Photorealistic Rendering . . . . .	239
<i>Jason L. Mitchell, Chris Brennan, Drew Card</i>	
Real-Time View Synthesis Using Commodity Graphics Hardware . . . . .	240
<i>Ruigang Yang, Greg Welch, Gary Bishop, Herman Towles</i>	
Recent Exact Aesthetics Applications . . . . .	241
<i>Tomás Staudek, Petr Machala</i>	
Reconstructing or inventing the past: A computer simulation of the unbuilt church by Alvar Aalto . . . . .	242
<i>Andrzej Zarzycki</i>	
Regeneration of Real Objects in the Real World . . . . .	243
<i>Hiroto Matsuoka, Akira Onozawa, Hidenori Sato, Hisao Nojima</i>	
Rich Curve Drawing . . . . .	244
<i>Suguru Saito, Akane Kani, Youngha Chang, Masayuki Nakajima</i>	
Sandwiching Surfaces . . . . .	245
<i>Jörg Peters, Xiaobin Wu</i>	
Shader Analytical Approximations for Terrain Animation in "The Time Machine" . . . . .	246
<i>Johnny Gibson</i>	
Shape Analogies . . . . .	247
<i>Aaron Hertzmann, Nuria Oliver, Brian Curless, Steven M. Seitz</i>	
Shape from Distortion: 3D Range Scanning of Mirroring Objects . . . . .	248
<i>Marco Tarini, Hendrik Lensch, Michael Goesele, Hans-Peter Seidel</i>	



## Sketches & Applications

Chair: Doug Roble (Digital Domain)

Shape-Based Character Animation <i>Eric Guaglione, Doug Sweetland</i>	249
Sketchy Rendering <i>John Haddon</i>	250
Slow Growing Volumetric Subdivision <i>Valerio Pascucci</i>	251
Star Fields in 2D <i>Maria Giannakouros</i>	252
“Still I Rise” Painterly Animation off the Shelf <i>Umesh Shukla</i>	253
Stylized Flowing Hair Controlled With NURBS Surfaces <i>Ramon Montoya-Vozmediano, Mark Hammel</i>	254
Synchronous Pronouncement <i>Sandra Villarreal</i>	255
Tangible Viewpoints: A Physical Interface for Exploring Character-Driven Narratives <i>Ali Mazalek, Glorianna Davenport, Hiroshi Ishii</i>	256
Techniques for Interactive Audience Participation <i>Dan Maynes-Aminzade, Randy Pausch, Steven M. Seitz</i>	257
Textures from Nonlinear Dynamical Cascades <i>David Mould, Eugene Fiume</i>	258
The AcceleGlove: a Whole-Hand Input Device for Virtual Reality <i>Jose L. Hernandez-Rebollar, Nicholas Kyriakopoulos, Robert W. Lindeman</i>	259
The AR-ENIGMA – A PDA based Interactive Illustration <i>Volker Paelke, Joerg Stoecklein, Lennart Groetzbach, Christian Geiger, Christian Reimann, Waldemar Rosenbach</i>	260
The Development of a Functional Visualization System for the Creation of Digital Human Models <i>John Jay Miller, Weidong Wang, Gavin Jenkins</i>	261
The Free-form Light Stage <i>Vincent Masselus, Philip Dutré, Frederik Anrys</i>	262
The Meditation Chamber : A Debriefing <i>A. Fleming Seay, Diane Gromala, Larry Hodges, Chris Shaw</i>	263
The Nail-Mounted Tactile Display for the behavior modeling <i>Hideyuki Ando, Takeshi Miki, Masahiko Inami, Taro Maeda</i>	264
The Spatial Bi-directional Reflectance Distribution Function <i>David K. McAllister, Anselmo A. Lastra, Benjamin P. Cloward, Wolfgang Heidrich</i>	265
The simulation of fluid-rigid body interaction <i>Tsunemi Takahashi, Heihachi Ueki, Atsushi Kunimatsu, Hiroko Fujii</i>	266
The Time Geography Project - using computer graphics to visualize problems in social science <i>Kajsa Ellegard, Johan Torne, Henric Joanson, Anders Ynnerman, Matthew Cooper, Mark Ollila</i>	267

## Sketches & Applications

Chair: Doug Roble (Digital Domain)

Three Angry Men: Dramatizing Point-of-View using Augmented Reality <i>Blair MacIntyre, Jay David Bolter, Jeannie Vaughan, Brendan Hannigan, Emmanuel Moreno, Markus Haas, Maribeth Gandy</i>	268
Towards Visualizing HCI for Immersive Environments: The Meta-Situational Tracker <i>Christopher Jaynes, Joan Mazur, Cindy Lio</i>	269
Tracking 3D Puzzle Pieces for Collaborative Learning Environments <i>Lori Scarlatos, Shalva Landy, Saira Qureshi</i>	270
TWISTER: A Media Booth <i>Kenji Tanaka, Junya Hayashi, Yutaka Kunita, Masahiko Inami, Taro Maeda, Susumu Tachi</i>	271
Uber Destruction in “The Time Machine” <i>Brad Parker</i>	272
User Customizable Real-Time Fur <i>John Isidoro, Jason L. Mitchell</i>	273
Using Cellular Phones to Interact with Virtual Environments <i>Bent Dalgaard Larsen, Jakob Andreas Bærentzen, Niels Jørgen Christensen</i>	274
Using the Virtual Terrain Project to Plan Real Cities: Alternative Futures for Hangzhou, China <i>Michael Flaxman</i>	275
Video Textures Using the Auto-Regressive Process <i>Neill Campbell, Colin Dalton, David Gibson, Barry Thomas</i>	276
Virtual Human Interfaces: Towards building an intelligent animated agent <i>Bernadette Kiss, Gábor Szijártó, Barnabás Takács</i>	277
Virtual Reality Interfaces Using Tweak <i>Patrick Hartling, Allen Bierbaum, Carolina Cruz-Neira</i>	278
Virtual View Generation by Linear Processing of Two Differently Focused Images <i>Akira Kubota, Kiyoharu Aizawa</i>	279
Wegzeit - the Geometry of Relative Distance <i>Dietmar Offenhuber</i>	280

### Special Sessions

Chair: Jackie White (California State University - Los Angeles)

Introduction	281
Committee	282
Yoda and Beyond: Creating the Digital Cast of Star Wars Episode II <i>Organizer: Dawn Yamada (Industrial Light + Magic)</i>	283
The Fate of Play: Game Industry Revolutionaries Speak Out <i>Moderator: Jason Della Rocca (International Game Developers Association)</i>	284
Spider-Man: Behind the Mask <i>Organizer: Mary Reardon (Sony Pictures Imageworks)</i>	285
What's Up Doc? A Fond Remembrance of Chuck Jones <i>Organizer: Jill Smolin (Independent)</i>	287

## Studio

Chair: Dan Collins (Arizona State University / PRISM)

Introduction .....	288
Committee .....	290
BLUsculpt™ .....	291
<i>Bill Brody, Glenn G. Chappell, Chris Hartman</i>	
Cinema 4D XL: Advanced 3D Software for Educators and Studio Professionals .....	292
<i>DaShawn L. Hall</i>	
Drawing Circle .....	293
<i>O. Makai Smith, James Stewart</i>	
Collaborative Frameworks: A Proposal for an Archive in the Studio .....	294
<i>Lily Diaz-Kommonen</i>	

## Web Graphics

Chair: Simon Allardice (clingfish.com)

Introduction .....	295
Committee .....	296
A case study in web3d film-making: Horses for Courses .....	297
<i>Michela Ledwidge</i>	
A Distributed Interactive Composition Tool .....	298
<i>Márcio Costa, Jonatas Manzolli, Dan Sharoni</i>	
Adaptive Solid Texturing for Web Graphics .....	299
<i>Bing-Yu Chen, Tomoyuki Nishita</i>	
Architecting a Distributed Dynamic Image Server for the Web .....	300
<i>Alain Chesnais, Tim Beck, Rudy Ziegler</i>	
Architectural Studios Online: The "Internet Studio Network" .....	301
<i>Alfredo Andia</i>	
Banja: Flash Programming System & Game Design .....	302
<i>Guillaume Clary</i>	
Color Aesthetics for Web Graphics Creation .....	303
<i>Lynda Weinman</i>	
Combining Procedural, Polygonal and Bitmap Representations using XML .....	304
<i>Anders Henrysson, Mark Ollila</i>	
Content Creation with Anark Studio .....	305
<i>Eric Morin</i>	
Creating and Implementing Photographic 3D on the web .....	306
<i>J. Paul Nykamp</i>	
Creating Interactive 3D Web Content using AXEL .....	307
<i>Teresa Lang</i>	
CT (City Tomography) .....	308
<i>Akira Wakita, Fumio Matsumoto</i>	
EVENTSCOPE: Discovering Mars with Internet-Based Virtual Environments .....	309
<i>Peter Coppin, Karl Fischer, Natalie Koch, Dana Martinelli, W. Ronald McCloskey, Michael Wagner</i>	
Flash MX Live: Real-time Video and Audio Delivery in Multi-user Environments .....	310
<i>Robert Reinhardt</i>	

## Web Graphics

Chair: Simon Allardice (clingfish.com)

Geartion: the web3D content for children .....	311
<i>Fusako Nishikubo, Manabu Tanaka, Shinta Ookino, Hiroshi Ogihara, Kazuhito Ezawa</i>	
Geometry Compression for ASCII Scenes .....	312
<i>Martin Isenburg</i>	
Integrating Multiple Narratives: <i>The Mirror That Changes</i> .....	313
<i>Annette Weintraub</i>	
Interactive 3D Characters for Web-based Learning and Accessibility .....	314
<i>Ed Sims, Dan Silvergate</i>	
<i>_knowscape</i> , a 3D multi-user experimental web browser .....	315
<i>Christian Babski, Stéphane Carion, Patrick Keller, Christopher Guignard</i>	
Multilingual Flash Applications .....	316
<i>Viswanath Parameswaran</i>	
Pseudo-3D Photo Collage .....	317
<i>Hiroya Tanaka, Masatoshi Arikawa, Ryosuke Shibasaki</i>	
Rhythm Engine .....	318
<i>Hidenori Watanave</i>	
Scalable Vector Graphics (SVG): The World Wide Web Consortium's Recommendation for High Quality Web Graphics .....	319
<i>Dean Jackson</i>	
Scalable Visualization of Super High Resolution 3D Images for Museum Archiving .....	320
<i>Anup Basu, I. Cheng, A. Mistri, D. Wolford</i>	
SMIL - An Introduction .....	321
<i>Philipp Hoschka</i>	
<i>Spoiral</i> : An online ad-lib mystery .....	322
<i>Kazuyuki Okada, Masa Inakage</i>	
The Bridge: An Environment for Collaborative Design Learning .....	323
<i>Yasuhiro Santo, Catherine Hu, Mamata Rao</i>	
The Evolution of Animation: Bedrock Revisited .....	324
<i>Sandro M. Corsaro</i>	
The Raisio Archaeology Archive: Using Design to Build Collaborations .....	325
<i>Lily Diaz-Kommonen</i>	
The Reality Cluster: Realtime Multimedia Communicatoin with Persistence .....	326
<i>Branden Hall, Samuel Wan</i>	
The Xj3D Browser: Community-Based 3D Software Development .....	327
<i>Alan D. Hudson, Justin Couch, Stephen N. Matsuba</i>	
TTT: A Web community tool mediated by friends .....	328
<i>Yuichiro Haraguchi, Sakura Toyabe, Masaru Murata</i>	
VisSheet Redux: Redesigning a Visualization Exploration Spreadsheet for the Web .....	329
<i>T. J. Jankun-Kelly, Kwan-Liu Ma</i>	
Wegzeit - The Geometry of Relative Distance .....	330
<i>Dietmar Offenhuber</i>	



# Community Outreach – Introduction

The Community Outreach Program introduced SIGGRAPH to the San Antonio community with a series of special programs in the months leading up to the conference. Its purpose was to generate and disseminate computer graphics information and ideas in San Antonio, and to establish a network of ideas that will continue after the conference.

Community Outreach presentations were targeted to educators and students, and featured distinguished computer graphics professionals from the fields of animation and interactivity. Speakers included Karen Sullivan (Ringling School of Art and Design), Dan Collins (Arizona State University), James Mohler (Purdue University), John McIntosh (School of the Visual Arts), Doug Roble (Digital Domain) and Simon Allardice (Stanford University and Lynda.com).

Students and faculty from more than 15 area high schools and 8 area colleges and universities attended the preconference workshops.

Tours of selected events at the conference are also planned for educators.

Nancy Wood, Chair

# Community Outreach – Committee

## Chair

Nancy Wood  
San Antonio College

## Committee

Phillip Nelson  
Newtek Corporation

Marla Hart  
Dynatouch Corporation

## Contributors

James Mohler  
Purdue University

Karen Sullivan  
Ringling School of Art and Design

Dan Collins  
Arizona State University

John McIntosh  
School of the Visual Arts

Doug Roble  
Digital Domain

# Educators Program – Introduction

educate

Pronunciation: ej' ū kāt'

Definition: to train or develop the knowledge, skill, mind or character of, esp. by formal schooling or study

The SIGGRAPH 2002 Educators Program highlights the processes, techniques and technologies that are critical for educating future pioneers, practitioners and visionaries in computer graphics and interactive techniques. Contributions include diverse content from post secondary and K-12 institutions, focusing on those who use computer graphics as a teaching tool as well as those who teach computer graphics as a branch of learning.

The 2002 Educators Program continues the highly successful forum presentation format, where moderators and attendees create a collaborative environment where important issues and problems facing educators are discussed. Unlike traditional venues, the forum is unique in that it allows all participants to draw upon the combined knowledge, experience and viewpoints of everyone in an open arena for communication. This year we have forums focusing on game development curricula and approaches, studio views on demo reels, “teaching gems” for computer science and art and design, and forums on creativity, developing a knowledge base for computer as well as how educators can inspire the “Renaissance” student in computer graphics. In addition to the forums scheduled throughout the week, the Educators Program also provides traditional workshops, papers and panels that are designed to round out the program and provide a full-week of activities for the computer graphics educator.

I would like to thank the members of the Educators Program Jury and the various reviewers who provided valuable feedback and insight in the selection of content for the program. It is this group that ultimately shapes the program into what it will be, bringing into reality the tangible form and format from the overarching vision for the program. My personal gratitude is also expressed to April Ramey for assisting with the myriad details of the program, as well as Stephen Spencer for coordinating the publication details. Finally, I wish to articulate my thanks to Tom Appolloni and Janet McAndless, my colleagues at Purdue University, and, most importantly, my wife Lisa.

James L. Mohler, Chair



# Educators Program – Committee

## Chair

James L. Mohler  
Purdue University

## Jury Members

Mark Bannatyne  
Purdue University

Dennis Bouvier  
University of Saint Louis

Colleen Case  
SchoolCraft College

John Finnegan  
Purdue University South Bend

Lew Hitchner  
California Polytechnic State University

Scott Meador  
Purdue University

Carlos Morales  
Purdue University

Jacquelyn Ford Morie  
USC Institute for Creative Technologies

## Reviewers

Marc Barr  
Kathy Barshatzky  
Judy Birchman  
Jack Bresenham  
Deborah Chew  
Cher Cornett  
Steve Cunningham  
Gerry Derksen  
Danial Dochev  
Prospero dos Santos  
Chad Dresbach  
Art Durinski  
Dena Eber  
Anthony Faiola  
Glenn Goldman  
Lilli Green  
Ugur Gudukbay  
Richard Kopp  
Joe Lohmar  
Mike McGrath  
Valerie Miller  
Mark Ollila  
Pierre Pepin  
Mario Rui Gomes  
Jaiaoying Shi  
Dennis Short  
Jill Smolin  
Beatriz Sousa Santos  
Diana Stewart  
Ken Walker  
Howard Whitston  
Stephen Wroble

# Forum: A Knowledge Base for the Computer Graphics Discipline

Gary R. Bertoline  
Purdue University  
grbertol@tech.purdue.edu

Cary Laxer  
Rose-Hulman Institute of Technology  
laxer@cs.rose-hulman.edu

## Abstract

Computer graphics is a powerful medium used to communicate information and knowledge. It is a discipline whose time has come. Until recently it was a mysterious specialty involving expensive display hardware, considerable computing resources, and specialized software. In the last few years, computer graphics has found its way into the mainstream of society, from entertainment, to engineering design, to the web, and virtually every industry. Much of this has been the result of spectacular improvements in the price and performance of computer graphics hardware and software. Interactive computer graphics is finding its way into nearly every discipline, industry, home, hospital, theatre, football stadium, automobile, appliance, and engineering office. Computer graphics is or will have an impact on nearly everything we do. The discipline of computer graphics is like a wide-open frontier where no matter which direction you move you will find more opportunities and undiscovered applications.

Computer graphics is the pictorial synthesis of real or imagined objects from their computer-based models. A related field is image processing which is the reconstruction of models of 2D and 3D objects from their pictures (Foley, et al, 1996). Traditionally, computer graphics is contained within the well-recognized disciplines of computer science, electrical and computer engineering, and art and design. The discipline of computer graphics has traditionally focused on developing new software algorithms (computer science) and hardware innovations (electrical and computer engineering) with an attempt to force art and design principles into the mix. Those types of developments will continue but the discipline is beginning to mature to a point where it is more than just about hardware and software developments and innovations. More attention must be given to the effective and novel use of the hardware and software developments in the context of graphics communication. This is a much more holistic approach to computer graphics as a discipline. This holistic approach is more focused on the end user and how these tools can be used in much more profound and revolutionary ways.

Obviously there have been many major and profound applications of computer graphics in its short history. However, the recent past of computer graphics is but a prelude of what is to come. Computer graphics can and will touch every human in a number of ways from entertainment to assisting in finding the cure of the most dreaded diseases. Computer graphics can and will have a profound effect on every type of business, industry, government, education, and the home. But it will take a very special type of education to prepare this next generation of computer graphics specialists.

Computer graphics is a powerful medium but only if combined with the principles of information design. The principles of information design are universal, like mathematics, and are not bound to unique features of a particular language or culture. The universal nature of the graphic language makes it a powerful tool in today's society where collecting, analyzing, and communicating all the available information is so important. Information becomes knowledge and knowledge can be commun-

icated more efficiently through computer graphics. Knowledge can become power and is the catalyst for stunning new developments in virtually every field.

The pure computer graphics discipline of the future will not be in computer science, art and design, technology, or electrical and computer engineering. The computer graphics discipline of the future will have its legacy in all these disciplines but will look to merge the software and hardware technology with the human communication process, which will result in novel ways of solving problems and disseminating information. As shown in Figure 1, computer graphics is the overlap between art, science and technology, and psychology.

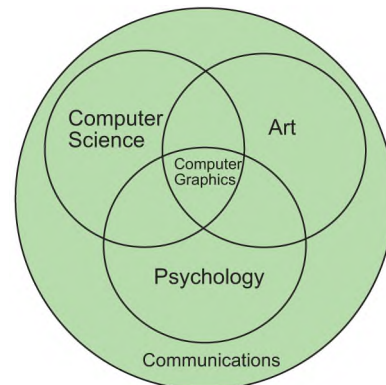


Figure 1 Computer graphics is the overlap between the art, science and technology, and psychology

The ideal student of this emerging discipline is bright, articulate, visual, analytic, and motivated by a passion for computer graphics. This student uses both sides of their brain but is keenly focused on the visual mode to solve problems. They are the modern-day Da Vincis, capable of visualizing what is nonexistent and finding solutions to complex problems. New opportunities are unfolding that require special talents and abilities for people with high visualization abilities who can use computer graphics tools to visualize scientific concepts and for the analysis and manipulation of complex three-dimensional information. As these new opportunities continue to unfold, the skill in manipulating and creating imagery may become more important than skill with words and numbers. Different kinds of tools may require different talents and favor a different type of discipline.

## References

- FOLEY, J. D., VAN DAM, A., FEINER, S. K. AND HUGHES, J. F. 1996. *Computer Graphics: Principles and Practice*. Addison-Wesley: Reading, MA.

# Forum: Game Development, Design and Analysis Curriculum

Jason Della Rocca  
IGDA  
jason@igda.org

Robin Hunike  
Northwestern University  
hunicke@cs.northwestern.edu

Warren Spector  
ION Storm Austin  
wspector@ionstorm.com

Eric Zimmerman  
gameLab  
eric@gmlb.com

## Abstract

Electronic gaming has changed. What was a curiosity twenty years ago is now one of the most popular forms of entertainment for all age groups. Games are now an accepted, pervasive part of US and world culture. The new ubiquity of games demands that we understand them as commercial products, aesthetic objects, learning contexts and cultural phenomena. This, in turn, creates a demand for people to fill a variety of roles associated with the creation, consumption and analysis of games.

For gaming to have a healthy future, industry and academia must find or create new contributors with new goals, ideas and experiences. We believe that educators have the best chance of creating and fostering these new goals and ideas while industry brings to the table a vital experiential understanding of the realities of game development. Academia and industry must work together, then, if progress is to be made.

Schools have begun to recognize the cultural and academic relevance of games. In addition to their increasing presence in the fabric of student life, games are rapidly evolving – diversifying and broadening their role as a communication medium. As universities design academic programs for the critical analysis of games and begin researching related interactive entertainment technologies, they see the value of including industry veterans in their plans.

The International Game Developers Association's Education Committee, comprised of both developers and academics, has been working on a curriculum framework and set of guidelines to aid in the development of game-oriented curriculum. During this session, the Education Committee's Curriculum Subcommittee will present its work to date, and solicit input and criticism on the framework. The goal is to generate significant feedback from the academic community to further refine the framework.

Here are some of the guiding principles used to develop the curriculum framework:

- ***The framework is not vocational but is career-oriented.*** The curriculum framework is focused not on churning out developers, per se, but on turning out well-rounded students of gaming, some of whom will no doubt choose to become developers.
- The IGDA's curriculum framework and related endeavors can be found online at the IGDA web site:
- <http://www.igda.org/Committees/education.htm>
- ***Gaming is an interdisciplinary field.*** This curriculum framework presents a wide range of topics encompassed by the acts of game creation, analysis and criticism. We strongly advocate cross-disciplinary study in any game-related course or curriculum.
  - ***Analysis, practice and context are equally important.*** Game development students must be exposed to analytical, practical and contextual materials.
  - ***The framework was designed to be versatile.*** Acknowledging the realities of educational institutions, especially when it comes to the introduction of new disciplines to the curriculum, we felt we had to create a game studies outline that was as versatile as possible.

# Forum: Inspiring the Renaissance Person

Donna Cox  
University of Illinois at Urbana-Champaign  
cox@ncsa.uiuc.edu

## Abstract

Originally, “university” was given its name because it was believed that a student would go to the university to find his or her place within the universe. Today's liberal arts education has limited this universal process and forces an early decision: the “major.” Students follow a well-defined curriculum that often precludes interdisciplinary study and cross-campus interaction. While most universities provide structure for in-depth, discipline-specific education; few offer curricula that take the student over many academic routes before making a decision on the major. Technical requirements can be the most constraining due to the vast foundational knowledge that must be mastered in the sciences and engineering.

Educational reward and economic systems force segregation of the disciplines. Faculty are often rewarded to be discipline specific, in both research and teaching. Students are not able to take classes across campus because the “major’s” students fill the courses leaving no room for non-majors. Universities are struggling due to low budgets and cutbacks. And it is very difficult to manage students who want to create new degrees that span campus boundaries.

In contrast to this segregation of disciplines, the computer graphics industry often demands generalist knowledge. Often a programmer is required to design or a designer is required to have advanced technical skills. It is apparent in the area of computer graphics, visual literacy is one of the most critical skills that span both art and science [West 1991]. A National Research Council committee studying these issues found a strain on the talent pool in 3-D computer graphics and that universities are failing to provide adequate cross-disciplinary training: “The development of workers with a mixture of technical and artistic capabilities represents a particular challenge because of its interdisciplinary nature. Whereas computer science and electrical engineering departments will train technical workers to address questions about networking and distributed simulation, the creation of visually literate workers demands cooperation between engineering and art departments, which are separated by large cultural and institutional gaps” [Zyda et al. 1997]. These cultural divisions have been talked about since C.P. Snow’s essay on “Two Cultures.” However there are many cultures and cultures within cultures. Within the educational system, there are many illogical divisions within and among departments. Certainly the divide between science and art is the greatest in education. How does this type of education prepare the student to find himself or herself in the universe? Given the constraints of contemporary educational systems, how can we educators inspire the Renaissance person, one who has truly interdisciplinary knowledge and skills?

Issues to consider:

- How do we focus students’ attention without constraining the breadth of knowledge?
- As educators, we are asked to prepare students for the “market”...do we really know the market?

- Should artists be taught to develop software as part of their curriculum?
- What are examples of organizing projects that encourage cross-disciplinary thinking?
- How does one find willing collaborative partners on large campuses?
- How can technology help in this quest to match collaborators and facilitate interdisciplinary activity?
- How can we change or by-pass organizational structures that inhibit the Renaissance person?
- Do collaborative teams in the classroom facilitate or hinder this kind of thinking?
- Given the vast knowledge that an interdisciplinary study must cover, should college terms be extended?

## References

- WEST, T., 1991. *In the Mind's Eye: Visual Thinkers, Gifted People with Learning Difficulties, Computer Images, and the Ironies of Creativity*. Prometheus Books, NY. p. 254.
- ZYDA, M., COX, D., KATZ, W., LARSON-MOGAL, J., LOUIE, G., LYPACZEWSKI, P., PAUSCH, R., SINGER, A., WESIMAN, J., 1997. *Modeling and Simulation Linking Entertainment and Defense*. National Academic Press: Washington, D.C. p. 93.

## Forum: Teaching Gems for Art and Design

Dena Elisabeth Eber  
Bowling Green State University  
deber@bgnet.bgsu.edu

Bonnie Mitchell  
Bowling Green State University  
bonniem@creativity.bgsu.edu

Heather Elliott  
Bowling Green State University  
helliott@hgnet.bgsu.edu

### Abstract

Despite the artistic maturity computer graphics has gained throughout the 1990s and into the new millennium, what we call the art student "wow" factor in computer graphics—the phenomenon of student obsession over new technology instead of artistic substance—is still a point of contention. It remains a challenge to teach computer graphics to art students while maintaining a balance between thoughtful art and sophisticated technology. In this session we will, with the help of the participants, reveal pedagogical solutions and uncover various approaches to teaching art students to use digital media in rich and expressive ways.

When the curve of new computer graphics technology was steeper and interfaces were not as intuitive as they now are, much of student learning was centered on the hardware and software and students struggled to comprehend digital media's place in the art world. In essence, we were in what Marshall McLuhan referred to as the first phase of a new technology—a stage in which, in the context of teaching digital media in the arts, the students were trying to understand computer graphics technology in terms of what they already knew. A digital image became a painting or a photograph, and 3D modeling and animation were understood in terms of film studies. Further, many students did not understand the scope of digital art and how it encompassed sculptural digital installations, how the art did not have to result in physical objects, or how the pieces could exist as interactive programs or websites. Students also worried more about the technology—what it could do and how to use it—rather than understanding the unique ways in which they could create aesthetically pleasing and profound works of art.

Today, computer graphics-based art is now closer to being established in the art world and embraces a more intuitive and sensible mode in which to create. However, the "wow" factor is still a problem and the changing pace of technology still keeps students and teachers alike tuned into the tool. Students are still trying to understand what the technology can do and often forget why or what they are making. Art instructors have a genuine struggle with balancing the time to teach students how to do something digitally with teaching them how to express themselves visually. Further, some software programs are truly complex and the concepts behind them require the student to understand, at least intuitively, the laws of physics, lighting, and how things change over time. Perhaps in some ways we are still in the first phase of digital technology in the arts.

From uncomplicated to elaborate software, how can instructors approach lessons without being bogged down with too much technology? Should instructors teach software or should they teach concepts, thus emphasizing the importance of technology as a means to an end rather than the end? What are some examples of lessons, assignments, or overall approaches that give proper weight to art and technology? How can teachers facilitate unique and expressive works of art?

Even if instructors are successful teaching individual expression with computer graphics technology, what domain within this vast field is most important? What are some areas on

which teachers could focus their energies? Are some computer graphics disciplines harder to teach than others? Is it more difficult to express ideas and compose formally beautiful works in some forms of digital graphics? Why or why not?

The best lessons and the best pedagogical approaches still will not hinder the dedicated special effects gurus from obsessing about what the software can do instead of the resulting work. How can instructors help these students care more about the art works?

Finally, and maybe most importantly, how can teachers keep up with the technology? Perhaps it is the teacher's concern with keeping current that influences the way they teach and what the students perceive as important. What are some solutions for instructors to keep up with changing computer graphics technology without compromising the core of the art they want to teach?

We ask the participants to come prepared to talk about some of these ideas and more. Once we target issues to resolve, we will ask the forum participants to put forward their gems: lessons, approaches, philosophical bends, and ideas that have worked or they believe will work. Attendees will leave the session with a notebook full of ideas and solutions to teaching students how to be expressive and create aesthetically pleasing works of computer graphics based art.

# Forum: The Role of Creativity in Computer Graphics Education

Bruce Wands  
School of Visual Arts  
brucewands@aol.com

## Abstract

This forum will give attendees a chance to present their own views on creativity and curriculum, as well as hear those of educators from a diverse group of colleges. Computer graphics education has grown tremendously in the last five years, particularly on the department level. Many issues have arisen related to the place of computer graphics education within a specific department's curriculum. They include the type of courses offered, challenges arising from the impact of this added educational component, and the desire to maintain traditional art education elements, such as theory and critique.

Creativity has historically been addressed in theory and critique classes. It is also now being taught along with software in computer graphics classes. Outcomes from this forum will allow educators to gain insight into their approach to nurturing creativity as it relates to computer graphics education. Other topics to be discussed include interdisciplinary approaches to curriculum, creating content for courses, and the relationship of computer graphics to traditional art education. The role of creativity in computer science classes, particularly programming, will also be discussed.

## 1 Background

Digital tools have become an integral part of the creative professions over the past decade. The Web is now a part of everyday life and an entire genre of computer animated feature films has evolved. Digital art received considerable public exposure in 2002 with major museum exhibitions at the San Francisco Museum of Modern Art, The Whitney Museum of American Art and the Brooklyn Museum of Art. Early pioneers in these fields were either artists interested in new technology or computer scientists interested in art. When the time came in the mid-1980<sup>TM</sup>s to start educating students to pursue careers in these emerging fields, educators were faced with dilemmas that continue to this day. Some schools teach art first before allowing the students to touch the computer, while others take a professional training approach and focus on teaching software. Institutional issues are still critical in defining where this education takes place. Does the Art Department buy computers or does the Computer Science Department teach art? Interdisciplinary concerns make solutions to these problems even more difficult. Designers are now working with print, the Web, animation, and video, in addition to design software. Many first year college students already have basic literacy in imaging, animation, video and design software. What major should they declare? Solutions to these issues have included bringing in outside professionals to lecture and critique, a renewed emphasis on traditional art skills, and the establishment of new degree programs.

## 2 Discussion Topics:

- Should individual departments add digital courses to their curricula, or should a separate department be created?
  - What are some of the revenue issues for colleges teaching computer graphics, based on current economic conditions?
  - What is the future of computer graphics education?
- 
- Exactly what is creativity and how do you teach it?
  - How do traditional art skills relate to computer art skills?



# Forum: Teaching Gems for Computer Science and Engineering

Mike Bailey

University of California San Diego  
mjb@sdsc.edu

Steve Cunningham

California State University, Stanislaus  
cunningham@siggraph.org

Lew Hitchner

California Polytechnic State University  
hitchner@csc.calpoly.edu

## Abstract

Moderators and attendees of this forum will present, discuss, and assess examples of "best practices" for teaching computer graphics to computer science and engineering students. Types of "teaching gems" to be presented include: classroom lectures and demonstrations, lab exercises, homework projects, self-instruction techniques such as tutorials, demo programs, and Web applets, use of analog models, examples from industry and the rest of the "real world", and field trips (real and virtual). The forum moderators will contribute presentations by electronic submission from virtual attendees prior to the conference and by attendees on site during the forum session. Three outcomes are anticipated from this forum: 1. presentation of several teaching gems, 2. group analysis and discussion of how best to apply each gem in a course and its learning effectiveness, and 3. posting of the results for public access on the SIGGRAPH Education Committee web site, <http://www.siggraph.org/education>.

## 1 Introduction

Teaching computer graphics to computer science and engineering students presents both a challenge and an opportunity. The challenge is that engineering related disciplines require learning about computer graphics in a way that helps students discover the trade-offs between visual realism and cost of computation. The opportunity is that students in computer science and engineering have the prerequisite knowledge to understand these tradeoffs. And, many students in these disciplines have math and science knowledge that allows instructors to use problems and applications in these domains for demonstrations and assignments that increase the student's motivation and effectiveness for learning about computer graphics.

However, students can waste a lot of time while learning graphics if they get sidetracked on tasks that are time consuming but not essential to learning, or if they get held up by having to deal with technical problems. For example, students are often tempted to spend lots of time developing complex 3D geometry models on assignments where the complexity of the model is unrelated to the main learning goals. Another example is assignments that require students to write their own computer programs. Students often spend a lot of time debugging program errors that are not related to learning goals for the project. However, these students often are willing to invest time on such tasks because they mistakenly believe the tasks are part of their required learning, because they find them fun or challenging, and because they are accustomed to having to deal with technical details.

Thus, when teaching computer graphics in these technical disciplines, there is a critical need for learning guided by an experienced teacher using proven teaching techniques. Many experienced faculty have discovered classroom and laboratory techniques that help students learn more rapidly and avoid spending excessive time on tasks that don't contribute significantly to learning goals. This forum will present a collection of "best practices" for teaching computer graphics to computer science and engineering students.

## 2 Definition of "Teaching Gems"

The phrase "Teaching Gems" for computer graphics is derived from the titles of the popular series of "Graphics Gems" books published by Morgan Kaufmann Publishers [Glassner 1990; Paeth 1995]. At the 1999 Computer Graphics and Visualization Education conference in Coimbra, Portugal (<http://www.siggraph.org/education/conferences/GVE99/>), one of this forum's moderators Mike Bailey coined the term "Teaching Gems" as a label for a publicly accessible collection and archive of teaching resources. Other education conferences have sessions devoted to presentation of effective and novel teaching techniques, for example the annual ACM SIGCSE (Special Interest Group on Computer Science Education) Conference sessions titled "Nifty Assignments" [Parlante et al. 2002]. This forum will provide a means for members of the graphics education community to present and contribute to a teaching resources archive.

The teaching resources presented during the forum will include: classroom lectures and demonstrations, lab exercises and homework projects, self-instruction techniques such as tutorials, demo programs, and Web applets use of analog models (such as toys), examples from industry and the rest of the "real world", field trips (real and virtual), and any other method or technique offered by a contributor.

Teaching gems presented during the forum will be gathered from three sources: the forum moderators' personal experience, contributions submitted electronically prior to the conference (following a call for contributions made in Spring 2002 via electronic mail lists and web site postings, and) contributions from audience members.

The moderators anticipate that interaction with and discussion by the forum audience will allow not just presentation of computer graphics teaching gems, but also explanation and justification of their teaching effectiveness as well as ways to assess the outcomes of their use. Following the SIGGRAPH 2002 Conference a summary of the teaching gems presented, the discussions at the forum, and supporting materials for applying the gems will be posted on the SIGGRAPH Education Committee web site, <http://www.siggraph.org/education>.

## References

- GLASSNER, A. 1990. *Graphics Gems I*. Morgan Kaufmann Publishers.
- PAETH, A. 1995. *Graphics Gems V*. Morgan Kaufmann Publishers.
- PARLENTE, N., ESTELL, J. K., GARCIA, D., LEVINE, D. B., REED, D., AND ZELENSKI, J. 2002. Nifty assignments. In *Proceedings of the 33rd SIGCSE Technical Symposium on Computer Science Education*, ACM Press, D. Knox, Ed., ACM, 319–320.

## Panel: K-12 and Industry Partnering

Darlene Wolfe  
Dr. Phillips High  
School

Jeff Scheetz  
The DAVE School

Roger Cotton  
The DAVE School

Timothy Comolli  
South Burlington  
High School

Chris Stapleton  
UCF Institute for  
Simulation and Training

### Abstract

Professional animators and digital effects artists and companies often work with high school teachers to explore relationships that can be developed and maintained between high schools and industry. The goals of this panel are to make more educators aware of the on-line and in person mentoring, sharing, and peer training that can occur in graphics, animation and related fields for 1) improving curricula and integrated technologies, 2) creating or furthering personal student interest or as a career vehicle and foundation, and 3) encouraging teachers to become involved in, and comfortable with, the new media.

When Josh Spector wrote "Studios drawn to CGI features" for the *Hollywood Reporter* [January 2, 2002], he noted the studios™ big interest in animation and computer graphics based on *Shrek*, *Monsters, Inc.* and the realization that there were audiences for the non-Disney animated feature. Youth of all ages, as well as adults, not only enjoy CGI, but can also create it themselves with the current technology on home computers.

Today's high school students, with the current technology available in home machines, are able to create products comparable to that of professionals, using computer technology that was very expensive just a few years ago. Although many students do not have access to the Internet or recent computer technology, those that do are producing programs for the Internet, their schools, and even, for sales. To gain and maintain the skill, one must practice. To give every child an even chance, the technology must be made available through education. Ideally, students would be given or lent computers with appropriate software for their classes, skills and academic or hobby interests.

In order to do that, administrators and teachers today must first recognize the convergence of electronic media, information technology, communication, and entertainment. Then plans must be made or programs adapted to further the educational needs and foundations of the student (at any age).

Therefore, it is imperative that teachers and administrators integrate 21st century technology into the curricula. However, we are sometimes afraid or hesitant to make changes we do not thoroughly understand, much less that are new to us. Many of today's educators are nearing retirement and there is a growing shortage of teachers. There is usually a shortage of funds in public education, too. Schools, therefore, often fall behind in the technology learning curve. New and exciting curriculum may be welcomed with open arms and with great desire but a lack of resources, combined with fear and trepidation by the very people who must learn and teach it, may negate its implementation.

### 1 Purpose and Premise

Where such integration is taking place, such as South Burlington High School in Vermont, exciting results are being seen. In many school districts, a lack of funds, fear of the technology, or a lack of knowledge prevent such improvements from being made. This panel addresses the continuum from school districts leading in technology to those not yet on the path. The professionals and the teachers will share their knowledge and experiences with the schools they are associated with as well as those they are aware

of. Where programs can be developed that includes staff training, peer tutoring, mentoring and business-education partnerships, the students are able to integrate their interests with required curricula. In such cases, software and hardware are frequently donated or sold at cost to the educational institutions for inclusion as educational resources. The use of students as peer teachers and mentors as both teachers and advisers to students and faculty enable great gains in the learning curve.

Many hardware and software manufacturers provide significant educational discounts to make their products available within the school system. Some will also give software in return for student evaluations and testimonials ? whether positive or not. Vendors realize that students and faculty will usually choose products with which they are familiar and comfortable. Some companies develop products to introduce students and faculty to the world of computer graphics. Other companies provide tutorials, videotape training, or manuals for self-paced learning. When educators and students can combine resources, products, and vehicles for education, a new world of technology opens to them.

### 2 The Panelists

**Dr. Darlene Wolfe**, [moderator] is the Television Production teacher at Dr. Phillips High School's Visual and Performing Arts School and has won two consecutive awards for connecting learning to earning. She was also a 2000 recipient of the Disney Teacherrific Awards.

Wolfe serves as a Valencia Community College [VCC] dual enrollment instructor with classes held at Florida's Orange County's (OCPS) "A"-rated Dr. Phillips High School TV studio and classrooms. She piloted the "2<sup>nd</sup> Unit" program with the Florida Motion Picture and Television Association (FMPTA). Traffic Productions is currently mentoring both 2<sup>nd</sup> Unit and M.E.D.I.A. (Mentoring Entrepreneurship by Direct Involvement in Activities) in a script-to-screen approach to a short film. In 20 years of teaching in Florida, Darlene has taught many subjects but gravitated to TV production due to the power of communication and electronic media in our society. She served on a SIGGRAPH panel in 1997 and has worked on SIGkids and Community Outreach at SIGGRAPH 1998. She has been fortunate enough to receive two SIGGRAPH Educators grants.

According to Wolfe, "Dr. Phillips High School is indebted to AIST and SCALA for their continued support and business partnership with our students. They have been most gracious and our students are learning on Iplay Studio, Info Channel Designer, Movie Pack, the earlier Media Mania, and now Movie 3-D. With these donated programs and the educationally priced LightWave 7.0 Educational Labpack, our students are able to learn with our mentors, as well as on their own."

With over 15 years in media production, **Jeff Scheetz** has amassed a thorough knowledge of visual storytelling through film, video and animation. He is a member of The Visual Effects Society and The Television Academy of Arts and Sciences. Prior to opening The DAVE School, Jeff worked for Foundation Imaging, NewTek

Inc., Nickelodeon Studios, Universal Studios Florida, Walt Disney Imagineering and The Disney MGM Studios. He and his wife founded the DAVE School (Digital animation and Visual Effects) located on the lot at Universal Studios Florida. Jeff not only serves as a guest speaker for high school students, he arranges tours and special classes to introduce and forward their interest in the digital world.

**Roger Cotton** is a professional 3D-character animator and instructor in San Diego, California. He graduated from the University of California at Los Angeles with a Bachelor degree in Theatre. He later attended the Musical Theatre M.F.A. Program at San Diego State University, studying concurrently at the Lehman Engal Musical Theatre Writing Workshop in Hollywood. Cotton worked as a multimedia content designer for five years before starting his 3D-animation career at Vision Scape Imaging, a 3D-production studio. He left to become a freelancer and was later appointed the Director of Multimedia at Platt College San Diego. There, he implemented and taught a twenty-week LightWave 3D training program. Cotton has created 3D art elements for various online game companies. He is currently a visiting professor, teaching 3D animation production at the DAVE School at Universal Studios in Orlando, Florida and consulting with nearby schools.

**Timothy Comolli** has been teaching at South Burlington High School in Vermont for 35 years. He came from a background in radio and television (he was the nation's youngest disc jockey in 1957) to teach Advertising, Video Production and Electronic Arts in the 90's. Because he was an English teacher by training, Tim's friends and colleagues find it hard to believe that he runs one of the best high school computer graphics facilities in New England. *Technology and Learning Magazine* and The Learning Company have named him National Technology Teacher of the Year. The Vermont State Legislature recently passed a resolution commending him for initiating the South Burlington Imaging Lab.

Tim has been awarded numerous other honors including: The Bill Gates "The Road Ahead" grant winner (1996), University of Vermont "Teacher of the Year" (1997), Who's Who Among Teachers (1997), The Henderson Foundation Grants winner (1999) and many others. Tim travels the country speaking to many technology conferences. You may reach him at [Tim@sburl.k12.vt.us](mailto:Tim@sburl.k12.vt.us). He has presented at SIGGRAPH 96 and was on a panel at SIGGRAPH 97. His students have achieved recognition and brought money and equipment into the "Imaging Lab." You can probably get the best pictures of the students work off their web page: <http://www.sburl.k12.vt.us>.

**Chris Stapleton** (see his web page for details <http://www.christopher.stapleton.net/>) is currently the Director for Entertainment Research at the Institute for Simulation and Training and Instructor of Interactive Entertainment for the UCF Film & Digital Media Program. His present work is developing content and infrastructure for the future of experience-based digital media for art, entertainment, and education. He is a producer of experience-based media by training. He has over twenty years experience as a creative principal in developing experiences and environments for entertainment, marketing and education for companies such as Universal Studios, Nickelodeon, Disney and Sanrio. His positions include Producer, Entertainment Consultant, Creative Director, Artistic Director, Design Director, Art Director, Chairperson, Concept Designer, as well as Production and Scenic Designer. Projects were developed for film, television, theater, theme parks, museums, research and education, retail merchandising, corporate marketing and communications. Experiences include projects worldwide including New York, Los Angeles, Tokyo, Milan, Cincinnati,

Knoxville and Orlando. His education consists of a Master of Fine Arts in Film and a Bachelor of Fine Arts in Theater from New York University, Manhattan and London. He has contributed to several publications including the book *Digital Illusions*, entertaining the future with hi-technology. He was recently the keynote speaker for the 2001 National Junior Scientist and Humanities Symposium; guest speaker at a 2001 Science Defense Board (US) Panel; moderator for the SIGGRAPH 2000 panel on simulation and entertainment, the Featured Speaker for the Fall 1999 SISO Simulation Interoperability Workshop; and keynote speaker for the 1999 international ACM Multi-Media Conference. He has been on the Executive Board of the Simulation Interoperability Standards Organization (SISO), as well on the conference planning committee for the IEEE VR conference, SIGGRAPH 94 and 98.



A sample of a Dr. Phillips High School student's virtual set



From a student at South Burlington High School



VR Research, Institute For Simulation and Training Industrial Partner, courtesy Chris Stapleton

# Panel: Animating Art History for Teaching

Roberta Tarbell  
Rutgers University  
RTarbell@aol.com

LiQin Tan  
Rutgers University  
Itan@crab.rutgers.edu

## Abstract

“Animating Art History for Teaching” is a creative learning methodology that incorporates new perspectives for introductory art history courses through digital technology created by computer animation and art history professors and their students.

**Keywords:** art history, 3D Animation, digital media, contrapposto, holy trinity, dome construction

## 1 Introduction

To overcome difficulties that students in introductory courses in art history have had in learning traditional concepts of art, the Professors of Art History and Computer Animation and our students in the Fine Arts Department have developed a new curriculum that synthesizes art history and animation. The approach uses animated spatial diagrams and sequences to teach several classical methods of the ancient Greeks and Romans that were revived by artists of the early Renaissance artists in Florence, Italy: one-point perspective in pictorial art, *contrapposto* in sculpture, and dome construction.

## 2 Exposition

“Animating Art History for Teaching” has devised a groundbreaking, collaborative methodology in animation for explaining difficult-to-understand concepts in art history. It provides an innovative and three-dimensional perspective towards teaching and learning art history through digital media. Students who have created it have learned art history with self-motivation, new interest and practical experience.

This unusual collaboration between educators in art history and computer animation catalyzed a synergistic effect and an unusually original product—the product is greater than a simple adding together of professional skills. Through collaboration, we discovered that principles central to the creation of sculpture, painting, and architecture as explained by art historians are the basic principles for animators. Z-Rhythm and Wave Principle, for example, refer to the anatomical structure and three-dimensional serpentine stance of *contrapposto*.

We also found it challenging to respect and harmonize the special qualities inherent in each of our fields—the accuracy, aesthetic connoisseurship and documentary evidence characteristic of art history and the advanced technology, humor, and anatomical and spatial exaggeration that distinguishes computer animation.

Digital animation is not only an expressive art vehicle but it also can give the artist tremendous freedom to create 3D history with striking impact. Art history stories vibrate with reminiscences of ethnicity and time through pixels and resolution by using a visual vocabulary of NURBS curves, polygons, clusters, particles, reflection maps, motion blurs and raytraced shadows.

## 2.1 Detail

From the onset, teams of professors and students collaboratively developed research, ideas, subject choices, storyboards, character designs, models, material textures, animations, renderings, and output. Throughout the process, professors actively listened to the unique insights of students and to professionals from very different fields of creative endeavor. We struggled to maintain a balance between the correctness of the art history content and its pedagogical mission and the caricature and comedy inherent in animation.

Students animated the *Holy Trinity*, a fresco in the Church of Santa Maria Novella, the masterwork of Masaccio (1401-28). Masaccio was a young and innovative artist in the fifteenth century akin to these student animators in the 21<sup>st</sup> century. When art history students observe the animated version of Masaccio’s *Holy Trinity* on a virtual turntable, they discover the three-dimensional constructs that Masaccio had in mind in 1425 as he conjured linear (orthogonal) perspective and foreshortening in this seminal fresco. Because Masaccio was breaking from a late Gothic (flat-style) of rendering the human figure and pictorial space, he had not fully mastered the mathematical and technical bases of his perspective pyramid that he centered below the cross on one point at the base. Our new video demonstrates this.

The *David* of Donatello (1386-1466) is as revolutionary as Masaccio’s frescoes. This freestanding, nude, bronze figure depicts the adolescent Biblical hero who symbolized Florence. *David* assumes a complex but stock pose of ancient sculptures called *contrapposto*, which had been devised by fifth-century B.C. Greek artists. After Donatello revived it, *contrapposto* dominated the pose artists chose for standing male figures in sculptures (including Michelangelo’s and Bernini’s later sculptures of David) and painted portraits. Trying to explain the three-dimensional serpentine line and the exquisite balance of human weights essential to this classic stance has challenged every art historian.

In 1410, when Filippo Brunelleschi (1377-1446) returned to his native Tuscany, he designed a dome to cover the 138-foot-wide hole of the octagonal drum of the *Duomo*, Florence’s Gothic-style Cathedral. Although Brunelleschi preferred the visual and geometric perfection of the hemispherical domes of the Romans, he had neither their technological experience nor their materials (concrete), and so, designed a steeper profile. He solved the problem of raising such a large dome with an ingenious method of horizontal bands that interlocked the eight vertical ribs at the corners of the octagon with sixteen smaller vertical ribs, and, also, at right angles, to the twenty-four ribs of the inner shell of the concentric dome parts. These two shells were stronger but weighed less than a single one.

These three pioneering early Renaissance artists and their masterworks were chosen because they were contemporaries who knew each other, worked synergistically with each other, and together, changed the face of European art for the ensuing

centuries. In some small way, we see ourselves collaborating to develop revolutionary methodology in teaching art history. Art history, a very young discipline in the humanities, which emerged in the curricula of a few universities during the early 20<sup>th</sup> century, is ripe for change in its presentation as well as in its content.

### 3 Conclusions and Future Work

The most dramatic and original contributions to 20<sup>th</sup> century art and culture involve digital media—not only in the creation of art, but also in the communication of art history. Forty years after 35mm color transparencies replaced 4" x 4" black-and-white glass slides as the standard for presenting images to art history classes and twenty years after videos were added to art historians' baskets of presentation tools for both lecture hall and seminar room, computer animation has the potential for revolutionizing the teaching of art history.

Several of the student animators will be present during this panel session to add their insights to the dialogue between presenters and conference participants. Although audience members will have their own queries, the presenters are interested in feedback from their peers and others who witnessed the presentation of this project. We are curious to know whether or not these didactic animations more effectively explain principles of art history than traditional methods of lectures, slides, and videos. How can computer animation be used for didactic purposes for art history and other fields of study? What are the optimal educational levels (K-12 or college) for adapting animated sequences on video or CD for teaching? How would you modify what we have achieved for a better product?

### References

- HARTT, F. 1993. *A History of Italian Renaissance Art*. 4<sup>th</sup> Ed. Prentice-Hall.
- MARAFFI, C. 2000. *Softimage XSI Character Animation f/x and Design*. Paperback
- ODOMS, L. S. 2001. *Art Across Time*. 2<sup>nd</sup> Ed. McGraw-Hill.
- ROSSANO, A. AND ARIMA S. 2000. *XSI Illuminated: Foundation 2*. Paperback
- STOKSTAD, M. 2002. *Art History*. Prentice-Hall, Inc.

# Problems with Using Components in Educational Software

Anne Morgan Spalter  
Brown University  
ams@cs.brown.edu

## Abstract

Reuse is vital in the education world because the time and money necessary to create high quality educational software is prohibitive. Estimates for the cost of creating a single well designed, highly graphical and interactive online course in the commercial domain range from several hundred thousand dollars to a million or more. Thus the idea of reusable software components that can be easily shared is tremendously appealing. In fact, “component” has become a buzzword in the educational software community, with millions of dollars from the National Science Foundation and other sponsors funding a wide variety of “component-based” projects. But few, if any, of these projects, have approached the grand vision of creating repositories of easy to reuse components for developers and educators. This paper investigates some of the factors that stand in the way of achieving this goal.

We begin by defining the word component and looking at several projects using components, with a focus on our Exploratories project at Brown University. We then discuss challenges in: Searching and Metadata, Quality Assurance, Programming in the University Environment, Platform and System Specificity, Social Issues, Intellectual Property Issues, and Critical Mass. We look at relevant software engineering issues and describe why we believe educational applications have unique factors that should be considered when using components.

Keywords: Components, education, educational software, reuse.

## 1 Why are Components of Such Concern to the Educational Community?

The concept of reusability for sharing elements of educational software is of special concern for the education community. Reuse is vital in the education world because the time and money necessary to create high quality educational software is prohibitive. Estimates for the cost of creating a single well designed, highly graphical and interactive online course in the commercial domain range from several hundred thousand dollars to a million or more. Thus the idea of reusable components that can be easily shared is tremendously appealing and few argue that reuse should be not be fostered at all levels.

Reusable components should be of particular interest to the computer graphics education community because one of the most frequent (if not the most frequent) types of educational components are graphical elements, from user interface widgets to visualizations and simulations. In the course of creating several dozen educational applets for the field of computer graphics, we have identified many areas that would benefit greatly from the existence of component building blocks, especially in the 3D domain where writing from scratch is especially difficult and time-consuming.

## 2 What, Exactly, is a Component?

Although there is virtually complete agreement about the end vision of reusability, the word “component” is subject to many different interpretations. The most frequent interpretation in the educational community seems to refer to any type of material that can be reused in an educational effort, from a lesson-plan to an applet to an image, video clip, or a piece of code. This meaning is in keeping with the dictionary definition of a component as simply “an ingredient.” and also seems synonymous with the vague term “learning object”, which is used to identify any part of a learning project or “environment.”

This all-encompassing definition of “component” obscures many issues, however. The problem of keeping track of a large collection of applets or media elements such as film clips is, on the technical side, chiefly a database and user interface problem. The problem of creating pieces of code that can be plugged together by programmers, however, is another sort of technical challenge that concerns a different audience. We are interested in the latter meaning and, in order to distinguish between our usage and the more general sense of the word, we will refer to these reusable pieces of code as “software components.”

All software components have certain high-level characteristics in common: they are pieces of code that provide documented, standardized means by which to access the functionality they contain. Popular standards include, among others, Microsoft’s Component Object Model (COM), Sun’s JavaBeans or Enterprise JavaBeans, and Borland’s Delphi VCLs.

Within this high-level definition, software components vary by how much a user can alter them. At one end of this spectrum lie “black box” components that are typically distributed in compiled or binary form and cannot easily be changed. Programmers have access to documentation about a black box component’s functionality and how that functionality can be accessed, but cannot change its code. At the other end of the spectrum lie “white box” components, which are typically distributed as source code and are therefore open to full inspection and change by programmers. Which end of the spectrum one favors is a matter of passionate debate among software engineers. Obviously, there are trade-offs with each approach: a black box component may be easier to use and test but may lack a key piece of functionality that the programmer desires. The white box approach lets programmers modify components easily but can easily lead to maintenance nightmares and make quality assurance more difficult.

In our Exploratories project we have adopted a “gray box” approach in which we make the internal structure of more complex components visible to programmers and permit them to replace small “black box” pieces or rewrite them as needed. This approach minimizes unforeseen internal dependencies that can occur in black-box component assembly while also avoiding some of the maintenance complications of full-fledged white-box components. More detail on our approach to component granularity can be found in E-slate [E-slate]. We recommend Clemens Szyperski’s Component Software book [Szyperski 1999]



for a more detail on technical and related social issues of software component use.

### 3 What Educational Software Projects are Developing Components?

A number of educational projects have been tackling the problem of creating reusable software components for educational applications, including our Exploratories project at Brown University [Exploratories Project]. Exploratories creates Web-based learning materials for the teaching of introductory computer graphics (currently using Java and Java3D) and publishes findings about strategies that seem to work for creating and using this content [Laleuf and Spalter 2001; Spalter et al 2000; Spalter and Simpson 2000]. In addition to our own experiences, we interviewed Chris DiGiano, a leader of the ESCOT project [ESCOT]. ESCOT created component-based software for middle school math. We also drew on email correspondence with Manolis Koutlis of the E-Slate project, a large-scale, long-term (over eight years now) undertaking in Greece to create an easy-to-use visual component assembly environment for teachers. Our understanding of both ESCOT and E-Slate were assisted immeasurably by reports from Andy diSessa's Web/comp project [WEB/comp Project], which is funded to explore issues in component-based educational computing. (Please note that the definition of component computing used in the Web/comp project is more inclusive than the one used in this paper.) DiSessa's write-up of his own project, Boxer [WEB/comp Project] sheds light on this subject. Finally, the Brown Exploratories project is part of the NSF National Science Digital Library (NSDL) program [NSF NSDL] and through annual meetings with fellow grantees we have become familiar with many issues of organizing repositories of educational materials.

We focus on the few projects mentioned above because these are the only large-scale educational software efforts we could discover that are attempting to produce sustainable libraries of educational software components. While there are many projects that assemble collections various types of stand-alone educational materials, such as applets (the Educational Object Economy, discussed by DiSessa is a good example), there seem to be very few that are building repositories of component-based building blocks in the sense that we define them in the previous section.

Despite some successes in these projects and others, and despite significant contributions from a range of talented and hardworking people, educators and programmers working with educators cannot yet go to the Web and search repositories of freely available reusable software components. We asked ourselves why all the work in components has not resulted in more of them and what can be done to fix the situation.

It is important to note that several projects, including E-Slate, as well as Exploratories (through collaboration with Dave Yaron's Irydium group [CREATE] at Carnegie Mellon) have also been working on the development of easy-to-use, predominantly visual environments that will allow non-programming educators to combine components on their own. E-Slate is currently offering free downloads of both its runtime and development environments [E-slate]. Boxer [WEB/comp Project], which lets educators create software in a custom environment that has many component-based aspects, also offers free downloads. The Carnegie Mellon tool is still under construction. These and other assembly tools being developed in the NSF NSDL should help bring teachers closer to the creative act of assembling their own educational software [NSF NSDL]. Developing such assembly tools is a research problem in and of itself and we do not investigate it further in this paper.

## 4 The Problems with Components

This section contains an alphabetical listing of central issues for component-based educational software projects and describes some of the problems others and we have faced or anticipate facing in the near future.

**Critical Mass:** The leverage provided by components cannot be realized without a critical mass of components. But the motivation to build component-based systems may be lacking if that critical mass does not already exist. This "chicken and egg" problem must be addressed for the vision of truly useful (i.e., high-quality and well-populated) repositories of educational components to become a reality.

There is ongoing debate about the number of components needed to establish a useful repository in a specific domain and the level of granularity that is most effective. ESCOT and E-Slate both favor the inclusion of large-grained, somewhat modifiable components (such as those created for ESCOT in Geometer's Sketchpad or E-Slate's Map, Database, or Grapher components). Exploratories, on the other hand, has focused on making all aspects of even large-grained components highly accessible to programmers, resulting in hierarchies of many fine-grained components. It may be that a repository designed for use by teachers who are combining pre-made components should contain a much smaller number of components (and chiefly larger-grained ones) than a repository designed for programmers who want to be able to make modifications at all levels (from what a simulation does down to its look and feel).

**Intellectual Property (IP) Issues:** The time and money necessary to create a working set of reusable components is most often prohibitive for not-for-profit ventures. Thus for educational institutions it can make sense to partner with commercial ventures or to try to market some aspects of one's work independently. The ESCOT project partnered with several companies that contributed greatly to the component content, particularly simulations. One of their main partners, AgentSheets, helped them generate simulations quite rapidly. Under the terms of the grant, many components were used for free for experimental educational work. Now that the grant is over, however, it leaves no free repository behind for others to use. Many IP issues have kept ESCOT from more broadly releasing their full set of components.

IP and licensing are standard issues in the corporate world and corporations have legal counsel and often their own team of lawyers. While university faculty commonly have some level of free legal counsel available for issues related to their work, teachers at K-12 schools are not accustomed to evaluating licensing options for code or paying for pieces of code and are most likely to avoid anything that is not offered for free.

**Platform and System Specificity:** Many projects, ours included, have run into enormously time-consuming problems addressing issues of platform specificity. This problem is not unique to education but is especially important because the types of computer platforms used in schools vary widely and many are relatively old, compared with corporate systems. We have given up trying to make our applets work on the Apple Macintosh series, for instance, because the Java plugins are routinely released months or even years behind those for the PC. Microsoft's .Net technology looks interesting, but won't run on Unix machines, which make up the bulk of the machines in our lab.

We have frequently encountered problems with different types of browsers (an applet works with one and not another) and among different versions of the same browser. In addition, the setup of the system software and networking for the machine (with attendant permissions issues) can cause problems. Both E-Slate and ESCOT have run into the same or similar issues in trying to roll out content into schools [Agesti and Evanco]. E-

Slate now uses its own software for both building and using its educational environments. Although these problems may seem trivial, they consume a disproportionate amount of time and can dramatically erode user confidence.

A partial solution to this problem, and one that has been adopted by our collaborators at Carnegie Mellon, is to package the software as well as the environment required to run it on a CD-ROM and physically distribute this to each student using the software. This lessens software incompatibility issues but does not resolve inter-platform or hardware compatibility problems. Such a solution also significantly lessens the impact of electronic learning since the course materials can then no longer be distributed electronically.

**Programming in the University Environment:** Unlike the full-time programmers in a corporation, the programming staff of an educational software project is often composed of undergraduate or graduate students. Although this approach is a rich educational experience for the students (and can also result in wonderfully creative ideas), we have found that the learning curve for component creation is much more difficult than we had imagined. Even after indoctrination into object-oriented programming and the benefits of reusability, and at least a month of motivating and then heavily using components (JavaBeans), we found a wide range of real understanding of what a component was and how one should be designed for reusability. Students were also unclear about how their components might interact with those from other sources. The E-Slate project has found that it takes about two years for an experienced developer to design truly reusable components.

In addition, even when students exhibit a reasonably strong theoretical understanding of how to design, write, and reuse components, and are able to adequately convey these ideas verbally and in writing, they are often unable to translate that knowledge into practical skills. Their software therefore seldom lives up to their own expectations or those of their collaborators.

Component projects undertaken in an educational setting, aside from coping with inexperienced programmers, are usually highly constrained by funding logistics as well. A grant may run for a period of only two, maybe three, years (E-Slate has been developed over the course of nine separate grants). Hiring talented staff without offering any sense of any job security can be difficult. In addition, it can often take many months for a new hire to come up to speed.

**Quality Assurance:** How do programmers know if the components they are downloading will work as promised? Will component source vendors fully test the wares they are selling? Few, if any, academic educational software projects employ any systematic testing of component compatibilities. Even testing by commercial component vendors such as ComponentSource [ComponentSource], a commercial Web-based marketplace for components, is composed only of testing for viruses, completeness, installation and de-installation. The projects discussed here are all in early stages of creating full-fledged component repositories and the original developers of the components are usually on hand to provide technical support and make changes. In the near future, however, resources will have to be directed toward testing efforts if the repositories of educational components are to be trusted (an essential feature for the whole component model to work).

Another approach to quality assurance lies in Amazon.com-style user reviews. ComponentSource offers user review sections and discussion groups for each of its components. The Merlot repository employs a peer review system (using teams of handpicked specialists) as well as user reviews for its assets; their framework could also be applied to software components.

**Searching and Metadata:** The repositories created as part of ESCOT's and E-Slate's work have so few entries that locating

items is not yet a serious issue (each has approximately 30 components). Components can be organized by categories that appeal to each project's programmers or organized by functional demands, such as different projects. For example, AgentSheets, a company that worked with the ESCOT project, provides an environment in which non-programmers can create graphical "characters" with behaviors. These characters can be exported as Java-Beans. Over 500 simple AgentSheets components are available on the AgentSheets web site [AgentSheets], organized by the name of the projects for which they were created.

While these strategies work with small numbers of components, different strategies are necessary for larger repositories. The two main approaches used in other types of reusable asset repositories will probably also work with repositories of software components: organization by topic and searching through metadata. A great deal of work on the manual and automatic generation of metadata has been done as part of the NSF Digital Library initiative. In particular, the education sub-program of this grant, the NSDL [Laleuf and Spalter 2001] has addressed issues particular to educational use of metadata and the environments in which it is used. Research has also shown that for domain-specific repositories of limited size, keyword searches without use of any metadata can be quite effective [Poulin 1999].

Several groups now use metadata to harvest content and to make it searchable. The Merlot project [Merlot], for instance, is a continually growing collection of online learning materials, peer and user reviews, and assignments. Such projects could extend their systems to include software components and we are currently in discussion with the Merlot group about this potential.

**Social Issues:** Finally, significant barriers exist in the sociological aspects of educational software component development. The ESCOT project often found that social and organizational issues had more impact on the results than any of the technical issues. The mentoring of both teachers (in the design and use of the software projects) and the software developers (in ways of working collaboratively with the teachers) often seemed to be a more important factor than any advances in the ease of coding.

Some of the problems crucial to understanding the use of components in K-12 education come from the need to combine strong software engineering skills, deep domain knowledge, and pedagogical or instructional design knowledge. The chief holders of these types of knowledge tend to have different backgrounds, personal goals, and terminologies. While our project at Brown largely circumvents such issues by having educators who not only talk the language of programming but actually teach it, the ESCOT project found first-hand that mixing these cultures can be a challenge. ESCOT used "integration teams," pairing programmers and designers with educators and educational technologists to design software that would have immediate classroom use. Although they confirm that having the classroom teachers in the design loop from the very beginning was essential, they acknowledge that communication can be difficult amongst the members of such groups.

## 5 Has Anyone Succeeded with Reusable Components in Any Domain?

Although the educational community has not yet had remarkable success stories with components, the picture is brighter in other fields. Large numbers of Microsoft Visual Basic components are available online. Marketplaces such as ComponentSource and Flashline [ComponentSource] have thousands of offerings and claim hundreds of thousands of downloads. Companies such as BEA [BEA], which sells component-based e-commerce software, provide proof of workable business models for the use of components in real-world situations.

Design and implementation strategies for creating component software have been developed at institutions such as Carnegie Mellon's Software Engineering Institute (SEI) [SEI]. The SEI group advocates a "product line" approach, in which reusable components are developed for specific domains according to strict software engineering guidelines. Such an approach can ensure that components offer robust functionality but are not so complex or overloaded with such as a wide range of functionality that programmers simply write new, simpler functionality for themselves. A coherent collection of components is much more likely to be worth learning how to use than a random collection.

Component-based reuse has also been judged a success in military projects and government institutions such as NASA [Agresti and Evanco 1992]. Although some fundamental research problems remain for general cases of component reuse and composition, many challenges have been met, including methods of domain analysis and ways of measuring the benefits of reuse [Poulin 1999].

Although there certainly have been successes, the SEI documents and virtually every other strategy or case-study paper reconfirm our own experiences that writing reusable code requires tremendous overhead. It may take many times as much effort to create a set of reusable components than simply to develop code for one-time use. In many cases it does not make sense to stress reuse. We have found that the final design of a reusable component must be completely solidified before undertaking this decidedly more demanding approach to programming.

## 6 So is This Just a Software Engineering Problem?

It is natural to ask why, if components are working in one domain, they aren't working in another? Is it simply that the proper software engineering knowledge has not been applied to educational uses or that the solutions to relevant problems are not even known? Our feeling is that the answer for the lack of success in the creation of reusable component building blocks for education cannot be ascribed solely to software engineering issues. Although better software engineering practices are factors and almost certainly are necessary, we do not think such knowledge and application of certain practices are, by themselves, sufficient to lift the educational component vision off the ground. Also, although the component problem is hard in any field, we do not believe that current limits of research knowledge are what are holding back educational efforts.

Our own experience and our discussions with others, in particular Chris DiGiano of the ESCOT project and Andy diSessa, reveal that production by and for the educational community differs substantially from that in more commercial settings. The social issues, platform problems, and severely limited funding and time-scales preclude the application of many software engineering recommendations, including, say, establishment of reuse teams in addition to regular programmers [Poulin 1999].

To add to these difficulties, there are a great many complex and poorly understood issues that affect educational software, making it difficult to analyze in terms of components. These issues include topics such as pedagogy, assessment, and classroom implementation. Even if a set of successful components is created for a given domain, the problems of interoperating with components from other subject areas, or within the same subject area but at different levels of sophistication and abstraction, are formidable.

## 7 Conclusion

A fully generalized vision of components, in which one can go to a Web site and grab many pieces of code, written by different

people for different domains, and easily assemble it into a application may never come to pass. This vision, may, in fact, be technically impossible. But a more limited goal, with domain-specific sets of components, written to standard specifications and tested before being released, still has the power to revolutionize educational software development, not only making it dramatically faster for programmers to create new applications but also making development accessible to non-programming educators through visual assembly environments. It is clear however, that the level of awareness of components needs to be vastly improved, both to motivate production and usage, and to help both project teams and funding agencies gain a realistic understanding of the needed resources. In addition, the vocabulary of components needs to be standardized so that we are all discussing the same things.

Even with increased knowledge and better software engineering, however, we do not believe that the essential critical mass of components can be created without some coordinated effort to "prime the pump." Although creating a "product line" of software components for all of computer graphics or some other large field may be impossible for many years, a coordinated effort to address, say, all the linear algebra needed for introductory computer graphics, or a color theory unit, could set an extremely useful example. However, such a collection, from user interface elements to underlying math engines or simulation pieces, cannot grow ad hoc; it would require a sizeable up-front planning and design effort.

Andries van Dam and others have proposed a Learning Federation, a non-profit consortium of government and industry, to provide an usually long-term and high level of funding for the creation of exemplary online courses using highly interactive computer-based environments. Something of this nature, with a focus on software components, may well need to be created before the potential of component repositories for educational software can be realized.

## 8 Acknowledgements

This paper would not have been possible without the technical insights of Shriram Krishnamurthi of Brown University and conversations with Chris DiGiano of the ESCOT project. We also wish to thank Manoulis Koutlis from E-Slate and Andy DiSessa and his Web/comp project. Thanks also to Andries van Dam for many close readings and to Jean Laleuf, for contributions to the technical sections, feedback on many drafts, and his contributing work as Technical Director of the Exploratories project. And, last but not least, a thank you to Rosemary Michelle Simpson for readings and the tracking down and organizing of the paper's references.

Our Exploratories work is sponsored by the NSF Science and Technology Center for Computer Graphics and Scientific Visualization and by the NSF NSDL grant "A Component Repository and Environment for Assembly of Teaching Environments (CREATE)." This work is also generously sponsored by Sun Microsystems, Inc.

## References

- AGENTSHEETS. <http://www.agentsheets.com/>.
- AGRESTI, W. AND EVANCO, W. 1992. Projecting Software Defects. In *Analyzing Ada Designs in IEEE Transactions on Software Engineering*, 18 [11], pp. 988-997.
- BEA SYSTEMS SUCCESS STORY. <http://www.componentsource.com/SellComponents/SuccessStories/BEASystems.asp>.

COMPONENTSOURCE. <http://www.componentsource.com/> and <http://www.componentsource.com/Welcome/NewVisitor/WhyBuyFromComponentSource.asp> for their testing statement.

CREATE (A COMPONENT REPOSITORY AND ENVIRONMENT FOR ASSEMBLY OF TEACHING ENVIRONMENTS) PROJECT. <http://www.smete.org/nsdl/projects/services.html#brwn>

E-SLATE. <http://e-slate.cti.gr/>.

ESCOT (EDUCATIONAL SOFTWARE COMPONENTS OF TOMORROW). <http://www.escot.org/>.

EXPLORATORIES PROJECT. <http://www.cs.brown.edu/exploratories>

LALEUF, J. R. AND SPALTER, A. M. 2001. A Component Repository for Learning Objects: A Progress Report. In *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries*.

MERLOT. <http://www.merlot.org/>

NSF NSDL (NATIONAL SCIENCE DIGITAL LIBRARY) PROGRAM. <http://www.smete.org/nsdl/projects/index.html>.

POULIN, J. 1999. Been There, Done That. In *CACM (Communications of the ACM)*, 42 [5].

SEI (SOFTWARE ENGINEERING INSTITUTE), Carnegie Mellon. A Framework for Software Product Line Practice - Version 3.0, <http://www.sei.cmu.edu/plp/framework.html>

SPALTER, A. M., LEGRAND, M., TAICHI, S., AND SIMPSON, R. M. 2000. Considering a Full Range of Teaching Techniques for Use in Interactive Educational Software: A Practical Guide and Brainstorming Session. In *Proceedings of IEEE Frontiers in Education '00*.

SPALTER, A. M., AND SIMPSON, R. M. 2000. Integrating Interactive Computer-Based Learning Experiences Into Established Curricula. In *Proceedings of ACM ITiCSE (Innovation and Technology in Computer Science Education) 2000*.

SZYPERSKI, C. 1999. *Component Software: Beyond Object-Oriented Programming*, Addison-Wesley, New York.

VAN DAM, A. 2000. Reflections on Next-Generation Educational Software. In *Enseigner L'Informatique: Melanges en Hommage 'a Bernard Levrat*. Pellegrini, C. and Jacquesson, A. editors, Georg Editeur, pp. 153-166. <http://www.cs.brown.edu/people/avd/LevratPaper.html>.

WEB/COMP PROJECT. diSessa, A. A. <http://dewey.soe.berkeley.edu/boxer/webcomp/index.html>.

# What Do Computers Eat? Teaching Beginners to Think Critically About Technology and Art

Tiffany Holmes  
School of the Art Institute of Chicago  
thome@artic.edu

## Abstract

This paper presents new curriculum for an introductory course in art and technology in which students compare the software industry with fast food to investigate patterns of consumption in our culture.

## 1 Introduction

Educators face a tremendous challenge in striking the proper balance between technique, craft, and theoretical knowledge in an introductory course in the media arts. What kind of course might orient students to the complicated history and cultural significance of both hardware and software while simultaneously teaching enough technological skills to allow them to create art works with substantive meaning and content?

In this paper, I will discuss an undergraduate course curriculum that I developed at the School of the Art Institute of Chicago. In “Fundamentals of Art With Technology,” students debate the similarities and differences between the fast food industry and the software business to gain a critical understanding of the forces that define patterns of consumption in our culture. In the following pages, I will describe the studio environment in which the course is taught and will provide details of reading assignments, class projects, and student evaluation.

## 2 Background

For the last six years, I have been teaching introductory and advanced classes in digital imaging and multimedia authoring at art schools, first in Baltimore, then Ann Arbor, and now Chicago. Of the two or three course syllabi I prepare each fall, the most difficult by far is the introductory class plan. At the School of the Art Institute of Chicago, this course offering is called “Fundamentals of Art With Technology” and is designed to orient students to the “concepts of electronic media, perception, inter-media composition, emerging venues, and other issues important to artists working with technologically-based media.” Every August, I anguish over which software packages to present to beginning students. Then I debate the merits of Flash versus Director and HTML versus Dreamweaver. I deliberate over readings: technical or theoretical emphasis? Have beginning students already read “Art in the Age of Mechanical Reproduction” in their art history classes? The questions are all too familiar to those faced with teaching an introductory course in the area of art and technology studies.

But the central question that challenges me each fall is ultimately far more consequential: how to introduce students successfully to the complicated history and cultural significance of the computer while providing enough technical instruction. The goal is to educate and empower students to create meaningful art works using tools unfamiliar and, in some cases, alienating to them. Computers are both medium and message, and in this duality lies the rub. No matter what the syllabus says, students prefer to think of the introductory class as a software-training

seminar. Instructors like me perceive the class to be a foundation on which to build critical thinking and problem-solving skills. The conflict of interest in the pedagogical focus inevitably produces unwanted friction between the instructor and students. Yet establishing the proper balance between technique, craft, and theoretical knowledge in “fundamentals” is critical to the department as a whole, because the atmosphere creates a ripple effect: the tone established in the beginning class often carries through to the upper level courses in the department.

This year, I had my most harmonious semester ever. I was thrilled to discover that students came to class informed and enthusiastic about the assigned reading. They arrived early to find out which new software package was to be placed under critical scrutiny, and they seemed eager to discuss the theoretical material and dive into hands-on software tutorials. What was the secret to the new rapport: fast food. My key to achieving group rapport and sustaining interest across all subject matter was to develop a curriculum that took a familiar subject’s history—the fast food industry—and compared it to the software business.

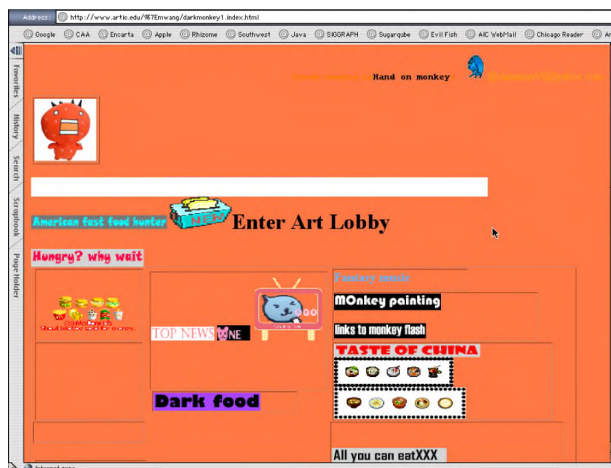


Figure 1 Personalized food portal, May Wang, 2001

## 3 First Class Meeting

We began our semester with a field trip to McDonald’s. Our guide led us through a whole world of gadgets designed to automate production of every menu item from the multi-layered Big Mac to the flawlessly blended McFlurry. Breathless with the excitement of a surprise outing, students compared notes about newly computerized cooking techniques—the sure-fire way to eliminate unwanted nasties like salmonella, E. coli, and other dangerous pathogens from the beloved burger. Back in the classroom, students divided into groups to react to the following general questions with the recent McDonald’s excursion in mind. What is the function of a menu? How does packaging affect perception of menu items? What strategies of production do fast food chains employ? What is brand loyalty and how is it created?

In the lively discussion that followed the brainstorming session, we quickly generated a correlation between the interfaces of the fast food industry and digital imaging software. For example, we compared the “Best Value” grouping of popular meal choices to groupings that software manufacturers use to establish default preferences, palette groupings, and unit settings. Products were more desirable when fresh—French fries become obsolete almost as quickly as software versions today. Brand name products seemed preferable to lesser known ones—most students admitted that if choosing from among Adobe Photoshop, Adobe PageMaker, and CorelDraw to add text to a photograph, they would select Photoshop since it was an acknowledged “brand-name.” Our laboratory activity for the afternoon involved an investigation of the Actions Palette in Photoshop and methods to automate the production of images in relation to one another. Further similarities to assembly line food production were noted after the demonstration of batch processing images for the web.

Readings assigned for the first class juxtaposed the history of the fast food industry with the history of Silicon Valley. In his book, *Fast Food Nation*, journalist Eric Schlosser dismantles the carefully constructed feel-good veneer of fast food and shows how the industry's astounding success has been achieved at astounding cost to the nation's health, environment, and culture. The book chronicles the events giving rise to the fast-food restaurant, beginning in 1948 when Richard and Maurice McDonald invented the Speedee Service System, introducing assembly-line efficiency into a commercial kitchen. In chapter 1, Schlosser describes the industry's pioneers—self-made entrepreneurs who pursued the American dream with old-fashioned ingenuity and a strong work ethic. Among these was Ray Kroc, who bought out the McDonald brothers and became the driving force behind the hamburger empire that is now the world's most recognizable brand name. Students found this journalist's text engaging, fact-packed, and straightforward. A selection from *The Silicon Boys* by David Kaplan complemented the Schlosser reading. Kaplan's book is a history of Silicon Valley since the 1930s, when Stanford professor Frederick Terman encouraged David Packard and Bill Hewlett to establish their own company in a garage. While the greed and excesses of venture capitalists prove distracting in later chapters, the initial ones provide an exciting account of the events that preceded the invention of the first personal computer and graphical user interface.

#### 4 Daily Schedule

The Fundamentals studio course met for six hours weekly. Prior to each class meeting, students read articles selected from two very different subject areas—generally one article focused on fast food and the other on the software industry or an artist's work. I required students to email me reading responses prior to class—this procedure worked remarkably well in that I was able to gauge individual performance while preparing discussion questions based on student contributions. Our daily schedule included three morning hours allotted for technical demonstrations and guided practice with lab tools. Ninety minutes of the afternoon hours were devoted to discussion of readings and peer writing. Guided questions encouraged students to draw their own comparisons between multinational corporations based on their inventive campaigns to capture and hold public taste: McDonald's and Microsoft, Taco Bell and Macromedia, the Mongolian Barbecue and Maya. Screenings of artist works and occasional collaborative class assignments occupied the final hour.

To give a specific example, the fourth class of the semester was set up in the following manner. Students read and responded to Schlosser's chapter entitled “Why McDonald's Fries Taste So Good.” In addition, they reviewed Simon Penny's article, “Consumer Culture and the Technological Imperative: The Artist

in Dataspace.” A media artist and theorist, Penny warns readers to approach seductive new technologies with a critical eye:

“The design of GUI's (graphical user interfaces), with their free choice among a fixed range of choices, is a mirror of the diner menu (ranch, thousand island, blue cheese, oil and vinegar) or the supermarket array. At the computer, as in the supermarket, one submits to the interactive scenario and the limited freedoms it offers: total freedom among a set of fixed options. A postmodern capitalist paradise!” [Penny 1995].

In this class devoted to comparing HTML web scripting to Dreamweaver's point and click interface, the Penny article provided an appropriate backdrop for students who struggle all semester to master the basics of tools like BBEdit and Dreamweaver as well as Photoshop and Director. Penny's description of a condition called “technofatigue” was particularly resonant with students. Throughout the semester, I work to select reading that will expand our dialogue during technical demonstrations beyond the basic queries regarding pull down menus and file sizes. Penny's essay is a must-read for all beginning students.

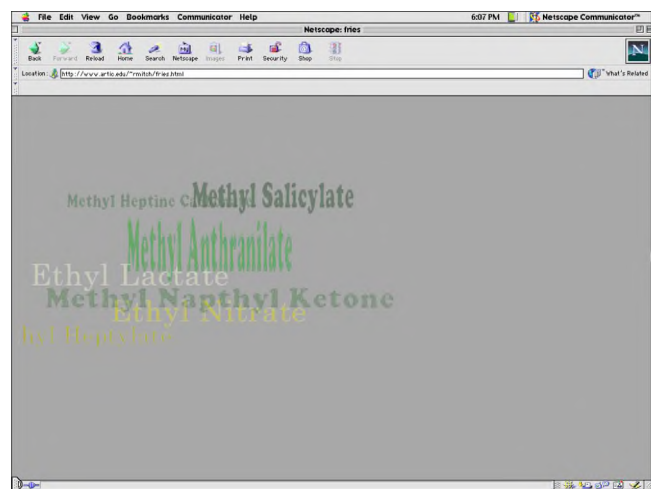


Figure 2 Animate composite, web site detailing French fry ingredients, Jeff Mitchell, 2001

For this fourth class, I designed technical demonstrations and practice assignments using Dreamweaver that would enhance our afternoon dialogue with the readings. In the morning session, I reviewed how to create web links using HTML code in BBEdit and then demonstrated how to create similar links in Dreamweaver. I also explain by example how to create a web page with links in Microsoft Word using the “Save as HTML” option. Students were asked to choose one software program to create a page devoted to the ingredients of French fries with multiple links to the manufacturers of each item. Many students created a text-based animation for the assignment, as shown in figure 2. In the discussion after lunch, we compared Penny's perception of the limited freedom allotted to each software applications' GUI with the factors that affect the flavor of French fries. According to Schlosser [2001], the taste of a French fry depends exclusively on the oil:

“For decades McDonald's cooked its French fries in a mixture of about seven percent cottonseed oil and 93 percent beef tallow. The mixture gave the fries their unique flavor — and more saturated beef fat per ounce than a McDonald's hamburger.”



Our conversation quickly led into the hazards of consuming processed foods that had hidden health risks or prepackaged meals with “secret recipes” like Colonel Sanders’ fried chicken. We noted the correlation between big-name manufacturers’ tendency to hide detail in the packaging and the composition of their product. Students compared the omissions inherent in ingredient lists boasting “natural flavors” with the Microsoft’s overly casual warning to the user that all margin and font settings would be changed during the conversion of a Word document to a web page. What you see is not necessarily what you get in a browser window or in a fast food restaurant.

After the discussion, I screened the Adbusters web site, the creation of a group of Canadian media artists who take an activist approach to cultural critique. The following quotation comes directly from their online mission statement:

“ADBUSTERS is dedicated to reinventing the outdated paradigms of our consumer culture and building a brave new understanding of living. We relish all truly political materials, whether they are scholarly probes into the decline of civilization, environmental forays into the forests, sci-fi carpet rides into cyberspace or humorous spoofs about commercial culture. More than anything, we seek compelling ideas that further the critical perspective and offer activist solutions. Our language is culture jamming: the new activism.”

Also featured on the Adbusters’ web site, the Ad Spoofs generated a good deal of interest and a meaningful lead-in to the afternoon studio hour. The collaborative assignment for the week was to create a “fast food spoof” web page. About half the class started to work to generate humorous images while half designed the web interface for our site. One individual began creating a banner ad for the site itself.

In general, students responded positively to the new curricular initiative although no one felt compelled to stop eating French fries. Final class evaluations indicated that the majority of students felt that the studio assignments, readings, and discussions relevant and meaningful to the course. Two students have continued to make artwork about patterns of consumption in our culture in advanced-level classes in the Art and Technology Department. For example, I am working with a student in my “Interactive Multimedia” course who is designing a breakout game designed to confound the user’s food preferences that were established previously in an online survey.

Most students indicated a desire on their evaluations to learn more in-depth skills to manipulate and control software. This outcome is very positive—the course was designed to entice a diverse group of students to focus their BFA study in our department. However, some students were discouraged that the class touched on several software applications instead of just one in depth. This kind of complaint seems to be common among students who hope to begin lucrative careers in web and multimedia design after graduation. As an educator, I believe that advanced software instruction can be easily accomplished with a manual or an online tutorial. From my standpoint, the critical ingredients of success in the class were the lively discussions provoked by the readings and screenings of artists’ work.

## 5 Conclusion

While I plan to rework some reading selections and study assignments, I will teach the Fundamentals class using the theme of fast food to spark interest and generate meaningful critique of the software industry and its products. Admittedly, the business of fast food is an arbitrary choice—interchangeable in some ways with other industries dependent on public taste and convenience. The comparison might not work as well in a different context. In

downtown Chicago, we are all but surrounded by the competing logos and combo deals offered by the restaurant chains desperate to lure us to lunch. The Big Mac will remain a favorite subject of mine to begin an important dialogue with aspiring media artists—the tools we use in the classroom come pre-packaged with all of the convenience, freshness and flavor that we demand from our burgers. How possibly, can we make art with software before we have some understanding of forces that generated that software’s popularity? That question can feed a class for a year.

## References

- BARDINI, T. 2000. *Bootstrapping Douglas Engelbart, Coevolution, and the Origins of Personal Computing*. Stanford, CA: Stanford University Press.
- KAPLAN, D. A. 2000. *The Silicon Boys and Their Valley of Dreams*. New York: Harper Perennial.
- LASEN, K. (editor). *Adbusters 'Zine*, Adbusters Media Foundation (<http://adbusters.org>).
- PENNY, S. 1995. Consumer Culture and the Technological Imperative: The Artist in Dataspace. *Critical Issues in Electronic Media*. SUNY Press. My students read an online version at: [http://www-art.cfa.cmu.edu/penny/texts/Artist\\_in\\_D'space.html](http://www-art.cfa.cmu.edu/penny/texts/Artist_in_D'space.html).
- SCHLOSSER, E. 2001. Why McDonald's Fries Taste So Good, excerpt from *Fast Food Nation*. Published in the *Atlantic Monthly*; January 2001; Volume 287, No. 1, pages 50-56. This piece is available online at: <http://www.mathaba.net/data/whymcdonalds.htm>.

# SIGGRAPH as Textbook: Learning Skills for Undergraduates

Kevin L. Novins  
University of Auckland  
novins@cs.auckland.ac.nz

## Abstract

This paper describes the design and implementation of a student-centered course in advanced undergraduate computer graphics. Instead of providing students with all necessary background before exposing them to recent research, students start with the latest SIGGRAPH proceedings and discover what topics they need to learn to understand them. The power of the group of students is exploited by having each student write a tutorial on one of the topics. Students then trade these written tutorials before being tested on the SIGGRAPH papers themselves. The course aims to nurture lifelong learning skills as well as an understanding of state of the art methods in computer graphics.

## 1 Introduction

The annual SIGGRAPH proceedings are a unique educational resource. Few other fields can boast such a compact collection of the most important papers of the year. As such, the SIGGRAPH proceedings are frequently a topic of study in advanced computer graphics seminar courses. However, the breadth and depth of background knowledge required for understanding papers in the SIGGRAPH proceedings pose a real challenge for educators. How can we give our students adequate preparation?

Simply expanding the range of prerequisite courses is not a viable option. Students have only so many electives to choose from and want to study cutting edge research as early as possible. Furthermore, the set of background topics required for the latest results in computer graphics is continually shifting.

This paper describes a student-centered approach to understanding the SIGGRAPH proceedings. Students start by choosing a set of the latest SIGGRAPH papers that the group will be responsible for. At first, they use the papers merely to identify gaps in their background knowledge. They then embark on individual research projects to fill those gaps. Each student works on filling only one gap, and then shares his or her newfound knowledge with the rest of the group. Ultimately, they are examined on the original set of SIGGRAPH papers.

This approach aims to increase confidence and to develop lifelong learning skills at the same time as it teaches computer graphics. It has elements of both peer assisted learning [Pigott et al. 1986; Wills et al. 1997; Donelan and Wallace 1998; Ploetzner et al. 1999] and problem-based learning. It was developed from earlier work in teaching computer vision [Novins and McCane 2001; Novins and Ming-Wong 2001].

## 2 Design and Implementation

The ideas outlined above were tested in a four-week unit of an undergraduate course in computer graphics at the University of Auckland. The earlier units of the course were taught in a traditional lecture-based format. Enrollment for the course was twenty fourth-year students with a half-semester third year course in computer graphics as a prerequisite. The stated goals for the student-centered unit of the course were:

- To identify active areas of computer graphics research.
- To develop research, writing and presentation skills.
- To work as a group to study one area of computer graphics research in depth.

The course had a scheduled group meeting three times a week for an hour. Each student's main task during the unit was to produce a written tutorial for use by the other students on an essential piece of background knowledge in computer graphics. However, there were a number of assessed activities leading up to the submission of the written tutorial. The course design and implementation are described in the following subsections.

### 2.1 Week 1: Selection of SIGGRAPH Papers

Since the group could not study all 65 papers from SIGGRAPH 2001 in depth in just four weeks, it was necessary to select a subset of them. We chose two papers from each of two sessions of the conference. The selection process was performed by the students in conjunction with a group activity designed to give everyone a brief overview of all 65 papers.

Each student was randomly assigned a session from the SIGGRAPH conference. Their job was to skim each of the papers in that session and make a quick evaluation about how interesting it would be to study and how difficult it would be to study. They were then to make a poster summarizing their findings based on a given template. Posters were assessed and counted towards 10% of the mark for the unit.

The first session of the week was an introduction to the unit and the second session was an optional question-and-answer session regarding the poster work. This left the third meeting for the poster session and a vote on which papers to study.

Students were enthusiastic about the poster project and all students did a good job. The poster session was lively, with students "campaigning" for their personal favorites.

At the end of the hour, we had three secret ballots. First to determine a short list of potential SIGGRAPH sessions, next to choose two sessions to study, and finally to select two papers from within each of the winning sessions.

The papers chosen were from the Hands and Words [Baxter et al. 2001; Cassell et al. 2001] and Animation and Expression [Sun and Metaxas 2001; Noh and Neumann 2001] sessions. After the session, five students were assigned to each paper.

### 2.2 Week 2: Choice of Topic Areas

The first two formal sessions of the second week were cancelled. Instead, each student had a half-hour individual meeting with the instructor to choose the topic for their tutorial. Before the meeting, each student was responsible for making an initial reading of their assigned paper, noting any confusing sections or unfamiliar jargon. Preparation for the meeting was worth 5% of the unit mark.



Students were given a handout identifying four criteria for a good tutorial topic:

- *Relevance* – Your topic should be chosen so that your final tutorial will help the reader to understand the methods, results, or context of your assigned paper in greater depth than they would simply by reading the assigned paper. Your topic need not be part of computer graphics itself.
- *Value* – The topic should be chosen so that the final tutorial will be of value to your peers. That is, it shouldn't be about something that most of them know about already ... or could easily learn with a quick internet search.
- *Research Potential* – The topic should lead you to research other sources besides your assigned paper. You'll need to find both library and Internet resources.
- *Uniqueness* – Your topic should be distinct from the topics of others in the class.

Discussion at the meetings focused on choosing a background area that met the criteria and identifying starting points for learning about that topic. It usually wasn't difficult to avoid overlapping topics, since different students had different backgrounds and interests. In a few cases, negotiation was necessary.

For the third session of the week, each student met with the others who were researching background knowledge for the same paper. This allowed them an opportunity to share insights on the paper and to plan a group paper presentation in the following week.

### 2.3 Week 3: Paper Presentations

During the third week, the formal sessions were devoted to short presentations of the papers by the students writing tutorials for them. These could not be comprehensive. They were meant to give all students an overview of the papers that they would be responsible for on the final exam. Students were encouraged to plan their talk around the following focus questions:

- What is the problem that the paper is trying to solve?
- What solutions are the authors proposing? How well do they work?
- How are the solutions different from existing solutions to the problem?
- What parts of the paper are most difficult to understand?
- How good are the results? Do they match the authors' claims?
- What are the limitations of the system?
- Are there any limitations that the authors don't report?

Each group was also responsible for asking questions of other groups after their presentations. They had to divide the tasks of formal presentation, question answering and question asking among themselves to spread the workload. Group performance was worth 10% of the unit mark.

The group presentation and associated activities helped to develop a healthy "team spirit" among those studying the same paper. The restriction of formal presentation to just fifteen

minutes and the requirement that every group ask at least one question of each of the other groups led to interesting discussion sessions. It encouraged broad understanding of the papers and critical analysis; however the sessions were light on technical details.

Another activity during the third week was a second individual half-hour meeting with the instructor. Students were expected to have made substantial progress on their tutorial research by this time and were asked to bring an outline and a list of references that they had uncovered in their research. Preparation for the second meeting counted for 5% of the unit mark.

### 2.4 Week 4: Peer Review

Students were required to write a complete draft of their tutorial and to bring two copies of it to the sessions in the fourth week. These drafts were distributed and each student read and reviewed as many drafts as possible. The review form contained the following instructions: "As reviewer, your task is to give the author of the tutorial constructive feedback about the content and clarity of the draft tutorial. Focus your comments on how the tutorial affects your learning. Be as specific as possible, and be sure to include both positive and negative feedback. Suggest ways that the tutorial could be improved." The draft tutorial and peer reviews counted for 20% of the unit mark.

Students took the draft tutorials and peer reviews seriously. It was an early chance for students to share knowledge and it enabled them to see how others were going about explaining complex material.

It also was an opportunity for students to discover whether they had pitched their tutorials to the right levels. While this succeeded in some cases where students had assumed too high a level of mathematical sophistication, most student reviewers were shy about giving negative feedback and about admitting that they didn't understand things. Many comments related to grammar, spelling and format rather than tutorial content. In retrospect, student reviewers should have been given clearer guidelines, perhaps including some sample reviews.

### 2.5 Final Tutorials

Final written tutorials were due at the end of the fourth week and were worth 50% of the unit mark. They were marked on both technical content and clarity. Students were told to use their SIGGRAPH paper as a general guide to length and format. Formal writing was not a requirement. Students were reminded that their audience was their peers and that they should avoid jargon.

As soon as they were received, final tutorials were put on the course web page so that other students could refer to them while studying for the final exam.

It was disappointing to discover that four of the twenty students engaged in plagiarism of published sources in their final tutorials. Three of those students were international students who may have had misconceptions about what was acceptable in addition to difficulties with written expression in English.

For the other students, final tutorials were of good quality, with a tighter spread of grades than the students' incoming grade point averages. This was perhaps due to the fact that they exchanged drafts as part of the peer review process and were able to increase or decrease their subsequent work in response to how they felt they were doing relative to the rest of the class. Weaker students performed better than they had in other courses and many of the top students performed worse.

## 2.6 Final Exam

A third of the closed-book two-hour final exam was devoted to the SIGGRAPH-oriented unit. Students were responsible for all four papers on the final exam. On the exam itself, students were given a choice of answering one of two essay questions. They were to support their answers using materials from at least three out of the four papers that were studied.

Performance on the exam correlated with performance in the unit assignments; however the overall quality was lower. This may indicate that the students did not find their peers' tutorials as useful as expected. However, there were several complicating factors including time pressure from other parts of the final exam, a short gap between handing in the tutorials and the final exam, and time pressure from other courses. More research is necessary.

## 3 Feedback

A student survey was performed just before the final tutorials were due. Fourteen out of the twenty students participated. They were asked to respond to the statements in Appendix A by rating their agreement as one of five steps on a scale from 0 (strongly disagree) to 4 (strongly agree). A mark of 2 is neutral. Average scores, standard deviations and departmental averages are given in Table 1. Departmental averages are across all undergraduate Computer Science courses.

In Questions 10 and 13, students rated the amount learned and the overall effectiveness of the teaching higher than the departmental average. In Question 1, all students agreed or strongly agreed with the statement "Material was presented in an interesting manner." In Question 12, just over three quarters of respondents agreed or strongly agreed with the statement "I would like more teaching in this format."

Students were also asked for written comments about the best and worst aspects of the course. In terms of the best aspects of the course, students cited the course content, individual choice of topic and accessibility of the instructor. One student commented: "[I] really enjoyed the change of style, and getting to study some of the newest developments in graphics. [It] was good to give students input into what was studied as it ultimately must increase interest. Often at university we take subjects thinking they should be really interesting and they focus on the boring stuff!" The aspect most cited as needing improvement was the workload. Although most students reported that the workload was fair, a substantial minority thought more time should have been given for researching and writing the final tutorial. One student noted that "consistent work [was] required ... [causing] strain on other assignments."

Statement	Mean	Std. Dev.	Dept. Mean
1. Presentation Interesting	3.50	0.50	2.51
2. Assignments Effective	3.14	0.52	2.76
3. Workload Fair	2.50	0.82	2.54
4. Lecturer Responsive	3.50	0.50	2.78
5. Lecturer Attitude	3.71	0.45	3.00
6. Lecturer Available	3.36	0.89	2.60
7. Diverse Backgrounds	3.00	0.82	N/A
8. Lecturer Enthusiasm	3.86	0.35	2.93
9. Lecturer Organized	3.00	0.38	2.79
10. Amount Learned	3.00	0.53	2.61
11. Interest Increased	3.00	0.78	N/A
12. Like Format	3.00	0.85	N/A
13. Overall Effectiveness	3.15	0.53	2.64

**Table 1: Results of the student survey. The numbers indicate average level of agreement with the statements, which are reproduced in full in Appendix A.**

## 4 Conclusions and Future Work

This paper presented the design of a student centered advanced undergraduate computer graphics course. Students start by selecting papers from the latest SIGGRAPH proceedings that they want to study, and work backwards through the literature to gain the background knowledge that they require. Through the use of shared written tutorials, each student is able to benefit from their peers' research, enabling them to achieve more in a limited period of time than they could on their own.

The course design was largely a success in the small study outlined here. Students were enthusiastic participants and developed computer graphics research skills at the same time that they learned about the latest techniques. They were pleased with the outcome.

The design is not without its flaws, however. An open question is how well the students were able to learn from each other's tutorials. This is an important area for future research. In the implementation outlined here, time pressure affected most of the students, probably reducing their ability to explore others' work. In the next incarnation of the course, the unit will be expanded to six weeks.

Time pressure affects the instructor as well. One-on-one meetings with students are an important part of the design, and require a large number of contact hours. This is balanced somewhat by a reduced load of lecture preparation. For larger classes, additional teaching assistants would surely be necessary.

Deep understanding and original tutorial content can be increased as a result of requiring a computer implementation as well as research and writing. This appears to be true even if the programming project is relatively small [Novins and McCane 2001; Novins and Ming-Wong 2001]. An interesting challenge will be to introduce implementation work into the problem.

Further work in this area seems well justified. As the joint IEEE/ACM task force on the Year 2001 Model Curricula for Computing (CC2001) states:

An essential requirement of any computer science degree is that it should enable graduates to cope with – and even benefit from – the rapid change that is a continuing feature of the computing field. But how does one achieve this goal in practice? At one level, the pace of change represents a challenge to academic staff who must continually update courses and equipment. At another level, however, it suggests a shift in pedagogy away from the transmission of specific material, which will quickly become dated, toward modes of instruction that encourage students to acquire knowledge and skills on their own.

Fundamentally, teaching students to cope with change requires instilling in those students an attitude of mind that promotes continued study throughout a career. [CC2001 2001].

This work is hopefully a step in the right direction.

## References

- BAXTER, B., SCHEIB, V., LIN, M. C., AND MANOCHA, D. 2001. Dab: Interactive haptic painting with 3d virtual brushes. In *Computer Graphics Proceedings, Annual Conference Series*.
- CASSELL, J., VILHJ? ALMSSON, H. H., AND BICKMORE, T. 2001. Beat: The behavior expression animation toolkit. In *Computer Graphics Proceedings, Annual Conference Series*.

- CC2001. 2001. Computing curricula 2001. Tech. rep., Computer Society of the Institute for Electrical and Electronic Engineers (IEEE-CS) and the Association for Computing Machinery (ACM), December. <http://www.computer.org/education/cc2001>.
- DONELAN, M., AND WALLACE, J. 1998. Peer assisted learning: A truly cooperative initiative. In *Students Supporting Students*, J. Dolan and A. Castely, Eds., SEDA Paper 105. Staff and Educational Development Association, 11–22. <http://www.ucl.ac.uk/epd/pal/SEDApaper1.html>.
- NOH, J., AND NEUMANN, U. 2001. Expression cloning. In *Computer Graphics Proceedings*, Annual Conference Series.
- NOVINS, K., AND MCCANE, B. 2001. Incorporating primary source material into the undergraduate computer vision curriculum. *International Journal of Pattern Recognition and Artificial Intelligence* 15, 5, 775–787.
- NOVINS, K., AND MING-WONG, S. L. 2001. A student centered approach to acquiring background knowledge in computer vision. In *Proceedings of the IEEE Workshop on Combined Research-Curriculum Development in Computer Vision*.
- PIGOTT, H. E., FANTUZZO, J. W., AND CLEMENT, P. W. 1986. The effects of reciprocal peer tutoring and group contingencies on the academic performance of elementary school children. *Journal of Applied Behavior and Analysis* 19, 93–98.
- PLOETZNER, R., DILLENBOURG, P., PRAIER, M., AND TRAUM, D. 1999. Learning by explaining to oneself and to others. In *Collaborative Learning: Cognitive and Computational Approaches*, P. Dillenbourg, Ed. Elsevier, Oxford, 103–121.
- SUN, H. C., AND METAXAS, D. N. 2001. Automating gait generation. In *Computer Graphics Proceedings*, Annual Conference Series.
- WILLS, C. E., DERMER, D., MCCAULEY, R. A., AND NULL, L. 1997. Studying the use of peer learning in the introductory computer science curriculum. Tech. Rep. WPI-CS-TR-97-11, Computer Science Department, Worcester Polytechnic Institute, September.

## A Student Survey Statements

Student levels of agreement with these statements are given in Table 1.

1. Material was presented in an interesting manner.
2. Assignments, tests, etc. were effective aids to learning.
3. Workload was fair.
4. Lecturer responded to students' questions in a constructive way.
5. Lecturer had a positive attitude toward students.
6. Lecturer was available to students outside class times.
7. Lecturer worked well with students from diverse backgrounds.
8. Lecturer showed enthusiasm for the subject.
9. Lecturer was well organized.
10. Amount learned overall (0=nothing, 4 = a lot).
11. My interest in the subject has increased as a consequence of this paper.
12. I would like more teaching in this format.
13. Overall effectiveness of teaching (0=very poor, 2=average, 4=very good).

# Mathematics And Geometry Education With Collaborative Augmented Reality

Hannes Kaufmann  
Vienna University of Technology  
kaufmann@ims.tuwien.ac.at

Dieter Schmalstieg  
Vienna University of Technology  
schmalstieg@ims.tuwien.ac.at

## Abstract

Construct3D is a three-dimensional geometric construction tool specifically designed for mathematics and geometry education. It is based on the mobile collaborative augmented reality system “*Studierstube*.” We describe our efforts in developing a system for the improvement of spatial abilities and maximization of transfer of learning. In order to support various teacher-student interaction scenarios we implemented flexible methods for context and user dependent rendering of parts of the construction. Together with hybrid hardware setups they allow the use of Construct3D in today’s classrooms and provide a test bed for future evaluations. Means of application and integration in mathematics and geometry education at the high school, as well as the university, level are being discussed. Anecdotal evidence supports our claim that Construct3D is easy to learn, encourages experimentation with geometric constructions, and improves spatial skills.

**Keywords:** augmented reality, mathematics education, geometry education, spatial intelligence

## 1 Motivation

Spatial abilities present an important component of human intelligence. The term spatial ability covers five components, spatial perception, spatial visualization, mental rotations, spatial relations and spatial orientation [Maier 1994]. Generally, the main goal of geometry education is to improve these spatial skills. In a long-term study by Gittler and Gluck [1998], the positive effects of geometry education on the improvement of spatial intelligence have been verified. Various other studies [Osberg 1997; Rizzo et al. 1998] conclude that spatial abilities can also be improved by virtual reality (VR) technology. However, little to no work has been done towards systematic development of VR applications for practical education purposes in this field.

To fill the gap of next-generation virtual reality interfaces for mathematics and geometry education, we are developing a three dimensional geometric construction tool called Construct3D [Kaufmann et al. 2000] that can be used in high school and university education. Our system uses Augmented Reality (AR) [Azuma 1997] to provide a natural setting for face-to-face collaboration of teachers and students. The main advantage of using AR is that students actually see three-dimensional objects, which they, until now, had to calculate and construct with traditional methods (mostly pen and paper). We speculate that by working directly in 3D space, complex spatial problems and spatial relationships can be comprehended better and faster than with traditional methods.

It is important to note that while at a first glance geometry education software appears to share many aspects with conventional 3D computer-aided design (CAD) software, its aims and goals are fundamentally different. Geometry education software is not intended for generating polished results, but puts an emphasis on the construction process itself. While relatively

simple geometric primitives and operations will suffice for the intended audience of age 10 to 20, the user interface must be both intuitive and instructive in terms of the provided visualizations and tools. Commercial CAD software offers an overwhelming variety of complex features and often has a steep learning curve. In contrast, geometry educators are interested in simple construction tools that expose the underlying process in a comprehensive way.

For productive use in the classroom, a number of circumstances must be accommodated: Support for a variety of social settings including students working alone and together, a teacher working with a student or teaching a whole class, a student or the whole class taking an exam, and so on. Collaboration in these situations is largely determined by roles, and the teacher should be able to retain control over the activities. Moreover, it is not realistic to expect that schools can afford extensive installations of expensive equipment such as is used in our lab, and therefore the software must run on a variety of immersive and non-immersive hardware platforms, including heterogeneous and hybrid setups.

This paper presents our current prototype of such an AR based geometry education tool, including hardware and software, user interface design, and initial experiences.

## 2 Related Work

Construct3D combines four research areas: geometry, pedagogy, psychology and augmented reality. We will give a short overview of closely related work in these areas.

A large body of work has been done on 3D modeling in general. However, although 3D input devices with six degrees of freedom (6DOF) have been used to enhance modelers, little modeling has been done in immersive virtual reality systems. A good overview of 3D modeling systems with 6DOF input devices can be found in the work of Mine [Mine 1996]. Mine’s CHIMP [Mine 1997] was used to study user interaction techniques. Similar goals were driving the work on DesignSpace by Chapin [Chapin et al. 1994], Bowman’s Conceptual Design Space [1996] and SeamlessDesign by Kiyokawa [2000].

A large number of researchers have been working on virtual reality applications for pure educational use [Bell and Fogler 1995; Bricken and Byrne 1995; Byrne 1996; Winn 1997 and many others]. A very good summary of existing educational applications is given by Mantovani [2001]. We want to point out CyberMath [Taxen and Naeve 2001], which is an avatar-based shared virtual environment aimed at improving mathematics education, combining the areas of virtual reality, pedagogy, and mathematics. Unlike our AR approach, CyberMath is intended for remote rather than face-to-face collaboration, and currently uses desktop VR with no support of immersive displays. In contrast, dedicated educational dynamic geometry desktop applications such as Geometer’s Sketchpad [Jackiw 1995], Cinderella [Richter-Gebert and Kortenkamp 1999], Euklid [Mechling 2000] and Cabri Geometry [Laborde and Bellemain 1998] support two-dimensional geometry only.

Regarding spatial intelligence, a recent article by Durlach et al. [2000] gives a very good overview of work that has already been done in the research area of enhancing spatial skills within virtual environments but mainly identifies the indispensable need for comprehensive future research in this area. Interesting work on gender differences has been done by Rizzo et al. [1998] and Larson et al. [1999]. Gittler and Gluck [1998] study how courses in descriptive geometry improve students' spatial intelligence. The main component of their results is that there is a significant positive effect of descriptive geometry instruction on performance in spatial ability tasks.

Finally, augmented reality is a rapidly evolving area of computer graphics and user interface research. A good overview is given in the survey by Azuma [1997] and its recent update by Azuma et al. [2001]. However, despite its obvious appeal for face-to-face collaboration, we are unaware of any AR application directly concerned with geometry education.

### 3 Application Design

The following sections describe two specific aspects of Construct3D: basic construction functions and layering.

#### 3.1 Basic Construction Functions

Construct3D is based on the *Studierstube* system recently described by Schmalstieg et al. [2002]. *Studierstube* uses augmented reality to allow multiple users to share a virtual space. We use see-through head mounted displays (HMDs) capable of overlaying computer-generated images onto the real world, thereby achieving a combination of virtual and real world and allowing natural communication among users. The latest version of *Studierstube* allows the mix and match of heterogeneous output devices such as personal HMD, virtual workbench, conventional monitors, and input through a variety of tracking devices. All these devices appear to act as interfaces to a single distributed system.

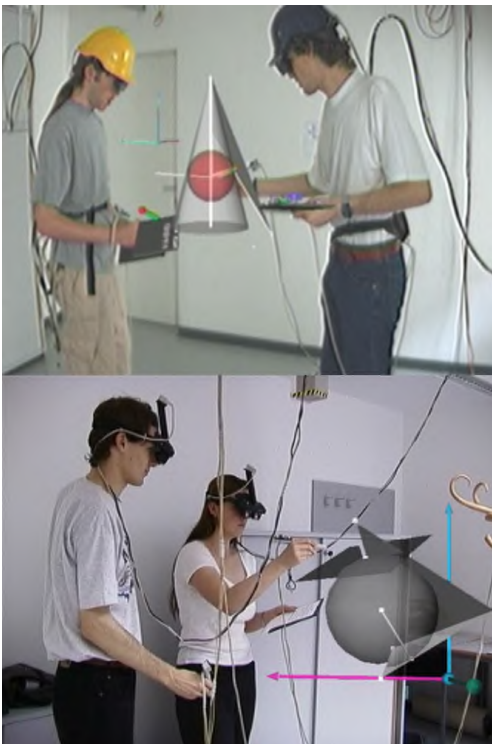


Figure 1: Students are working with Construct3D in our standard lab setup.

The current version of Construct3D offers a basic set of functions for the construction of primitives such as points, lines, planes, cubes, spheres, cylinders and cones. Construction functions include intersections, normal lines and planes, symmetry operations, and taking measurements. Recently Boolean operations based on the OpenCascade tool [2001] have been added that (for instance) enable learning about intersection curves of 2<sup>nd</sup> order surfaces.

Construct3D promotes and supports exploratory behavior through dynamic geometry (i.e., all geometric entities can be continuously modified by the user, and dependent entities retain their geometric relationships). For example, moving a point lying on a sphere results in the change of the sphere's radius.

All construction steps are carried out via direct manipulation in 3D using a stylus tracked with six degrees of freedom. AR affords that users see their own body and hand as well as the effects of their actions while working, so the construction process physically involves the students and resembles handcraft more than traditional computer operation. We believe that this is a key factor in the potential success of using AR for teaching geometry.

Necessary system operations such as selection of primitive type, load, delete, undo, and so on are mapped to a hand-held tracked panel, called the personal interaction panel (PIP) [Szalavari and Gervautz 1997]. The PIP allows the straightforward integration of conventional 2D interface elements like buttons, sliders and dials, as well as novel 3D interaction widgets. The haptic feedback from the physical props guides the user when interacting with the PIP, while the overlaid graphics allow the props to be used as multi-functional tools.

#### 3.2 Layers

We are using a 3D-layer system very similar to the one used in current image editing programs. Our 3D-layers offer the possibility to arrange parts of a construction into overlapping sub-spaces that can be controlled independently. This feature is particularly powerful in conjunction with distributed multi-user operation, where every user has a personal display for which visibility of layers can be controlled independently.

The current implementation allows the arbitrary selection of visibility per user and per layer. We have implemented three basic modes:

- *independent mode*, where every student can only see the elements constructed by himself
- *collaborative mode*, where everything is visible to everybody
- *teacher mode*, a special user ? the teacher ? can set visibility with a user/layer matrix of controls on the PIP.

Consider a teacher working on a construction with students watching him and work on the model themselves by request. The whole construction is visible to all users. If later the teacher wants the students to practice on the same construction again, he switches to "independent" mode while the application is still running. Now each student can only see the immutable specification and the elements that he constructed himself without being influenced by the work of the teacher or fellow students. If needed, the teacher is able to switch on his own construction or a reference solution so that some or all students can see it. The full solution to a construction task can be available from the beginning for reference purposes in a separate set of layers and can be progressively revealed by the teacher.

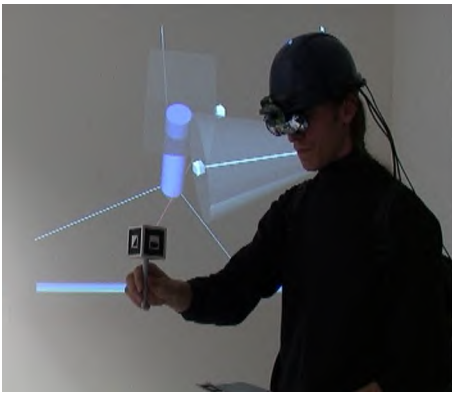


## 4 Hybrid Hardware Setups

To complement the diverse teacher-student interaction scenarios that are possible on the software side with practical hardware solutions for an educational environment we created various hybrid hardware setups. Realistically not all scenarios can be done in schools with equipment similar to our standard lab equipment of rather expensive tracking systems, HMDs and stereoscopic video projections. However, many components such as PC workstations with accelerated graphics and inexpensive projection systems are becoming feasible for classroom use. We are evaluating the following hardware setups:

**Hybrid AR classroom:** In our standard *Studierstube* setup, we have 2-3 collaborating users wearing HMDs for a shared virtual space, and holding fully tracked interaction props in their hands, as shown in Figure 1. In the top image in figure 1, students inscribe a sphere in a cone, while the bottom image portrays a simple example from vector algebra. The images were generated as live video capture with computer overlays.

While a hybrid AR classroom allows for first-class experiences on the student's side, the number of available AR sets significantly restricts use in larger groups. This situation is somewhat analogous to the use of a blackboard in class: Either the teacher or a single selected student work on the blackboard, while the remainder of the class watches or works along on paper. During a lesson, students take turns at the blackboard. With the aid of an additional computer with video camera and video projection screen, we can mimic this classroom procedure by projecting a live (monoscopic) video of the users (teacher and student) augmented with their current construction on a projection screen next to the users for the remainder of the class to watch (see Figure 2). Just as in conventional classrooms, students can take turns at using the HMD and working in front of the class.



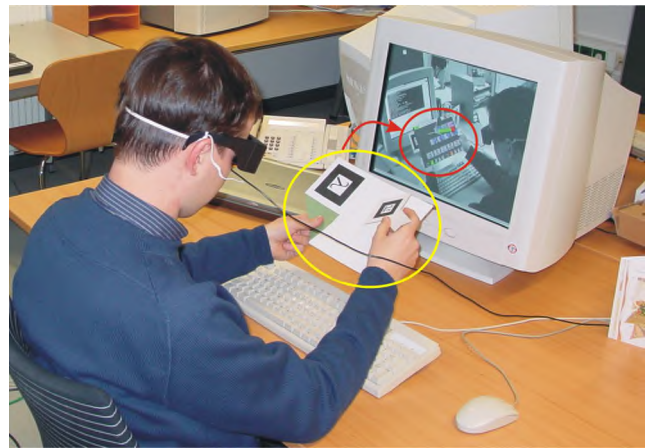
**Figure 2** A teacher is working in Construct3D with the mobile AR setup while a live monoscopic video of his current construction is projected onto a projection screen behind him.

**Projection screen classroom:** A popular semi-immersive technique is to use a single large screen projection shared by a group of users (in our case, the class), typically showing stereoscopic images using active or passive stereo glasses. The disadvantage is that since the screen is shared between the active user (e. g., teacher, demonstrator) and the observers, head-tracking is not useful. And, consequently stereoscopic images are often severely distorted if rendered for an “averaged” viewpoint. In consequence, manipulation even with tracked input devices becomes indirect (comparable to screen and mouse manipulation) as objects do not appear aligned or superimposed with the users hands. Advantages of this approach include lower system complexity and cost, as well as the avoidance of cumbersome HMDs. Despite the shortcomings, projection walls are established

techniques for semi-immersive group environments and single-projector displays are affordable for classroom use.

**Distributed hybrid classroom:** Just like the hybrid AR classroom, this setup may use personal HMDs for realizing AR for the teacher and selected students. However, the students are all equipped with personal workstations displaying desktop VR, watching the construction process on their screen.

We built a desktop VR system using a FireWire camera for optical tracking and a standard consumer graphics card with shutter glasses to get stereo rendering with optically tracked 6DOF input devices at a very low price. Figure 3 shows this system in use. A FireWire camera (out of view) is used for optical tracking of the hand held props that are equipped with markers (white ellipse). The image of the camera is used as a video background. Stereoscopic images are displayed on the monitor, which gives the user who wears shutter glasses the impression of working in 3D space. The virtual images of pen and PIP can be seen on the monitor (black ellipse) as an overlay over the video image.



**Figure 3:** A user working with the desktop VR system.

The advantage of this scenario lies in the relatively low price for a personalized semi-immersive display: Students can choose individual viewpoints; maybe even manipulate local copies of the constructed object. However, a teacher can also choose a guided mode, (e.g., by locking the students' views to the teacher's viewpoint).

**Remote collaboration:** Although the advantages of co-located collaboration are lost, the same systems can be used for remote collaboration through a remotely shared 3D space. For example, a teacher can remotely advise a student at a homework problem by the same guided construction techniques as in the AR-classroom scenario, or multiple students can remotely work together. Each of the users has an individual choice of input and output facility, (e.g., one user may wear a HMD, while another one uses a desktop VR setup). We are currently evaluating this possibility using our latest mobile AR hardware prototype [Reitmayr and Schmalstieg 2001] as a test platform.

## 5 Evaluations

The key hypothesis—that actually seeing things in 3D and interacting with them can enhance a student's understanding of three-dimensional geometry—were supported by the anecdotal evidence we have gathered from trial runs with real students. In our first evaluation [Kaufmann et al. 2000] with 14 students we got very positive and encouraging results and some problems were also pointed out. During the evaluation it was gratifying for

us to see users work with Construct3D in a very constructive manner. It was obvious that they did not need a long introduction to the system but applied their experience with 2D user interfaces to the 3D interface. The students' interactions with the system in our HMD-based lab setup were interesting to watch. After completing their task, some walked around the objects, viewing them from different sides or got down on their knees and looked at the scene from below. It was clear that they were proud of what they had "built." Half of the students felt that working with Construct3D for the first time is easier than a first experience with a desktop CAD package. All except one could imagine working with Construct3D without having worked with a traditional CAD package before.

Hand-eye coordination is very difficult when spotting a point accurately in 3D space without haptic feedback or constraints. All students reported problems with setting points at given coordinates. As a consequence we implemented raster and grid functions. About constructing in VR, students especially liked walking around and inside objects, the "playful" method of construction, and that spatial relationships and complex three-dimensional situations are directly visible. The clearness of Construct3D's menu system and the audio help system were mentioned positively, too.

Students mentioned the following possible application areas: interactive conic sections, vector analysis, enhancing spatial abilities, intersection problems, experiencing space (for very young students) and building three dimensional worlds from two dimensional views, elementary geometry, visualization of constructions, geometry didactics - learning by doing and training of spatial abilities by viewing objects from different sides.

At this stage students do not use Construct3D on a regular basis in mathematics and geometry education, but we plan to do extensive evaluations in current and upcoming research projects where students will actually learn by using our application. While developing Construct3D we are regularly visited by teachers, students, colleagues and friends who evaluate the system and give feedback on its quality. This helps to constantly improve the application and adopt it to the students' needs.

## 6 Conclusions and Future Work

In this paper we present a fully functional educational AR application for mathematics and geometry education. We implemented flexible methods to support various teacher-student interaction scenarios. Hybrid hardware setups allow the use of Construct3D in today's classrooms and provide a test bed for these scenarios. Initial evaluations of the concept are encouraging and our mid- to long-term plans are to integrate it in Austrian high school and higher education curricula. We have established partnerships with Austrian schools and scientific partners such as the Institute of Geometry at Vienna University of Technology and the Institute of Psychology at the University of Vienna.

Much work remains to be done. In particular, a comprehensive evaluation of the practical value of an education tool such as ours will require the development of substantial educational content that is put to real use in classroom. We are currently at the stage where we have working tools available, but now need to apply them to real educational work. For the beginning, we plan to create tutorials for vector algebra, conic sections and Boolean operations. We believe that despite the exciting possibilities of the new media, educational content creation for an interactive system is at least as difficult as authoring good textbooks and will require a substantial amount of time and work. Finally, the true value of the new tool in classroom use needs to be verified through controlled evaluations.

## Acknowledgements

The authors would like to thank Gerhard Reitmayr for his ongoing support, Klaus Feiler and Reinhard Steiner for integrating Boolean operations, and all of the Studierstube team. We also thank all participants in our pilot studies for their time. This work is sponsored in part by the Austrian Science Fund FWF under contract number P14470-INF and the EC under project number IST-2001-34204 (Lab@Future).

## References

- AZUMA, R. 1997. A Survey of Augmented Reality. *PRESENCE: Teleoperators and Virtual Environments*, Vol. 6, No. 4, pp. 355-385.
- AZUMA, R., BAILLOT, Y., BEHRINGER, R., FEINER, S., JULIER, S., AND MACINTYRE, B. 2001. Recent Advances in Augmented Reality. *IEEE Computer Graphics & Applications*, Vol. 21, No. 6, pp. 34-47.
- BELL, J. T., AND FOGLER, H. S. 1995. The Investigation and Application of Virtual Reality as an Educational Tool. In *Proceedings of the American Society for Engineering Education 1995 Annual Conference*, Session number 2513, Anaheim, CA.
- BOWMAN, D. A. 1996. Conceptual Design Space - Beyond Walk-through to Immersive Design. In Bertol, D., *Designing Digital Space*, John Wiley & Sons, New York.
- BRICKEN, M., AND BYRNE, C. 1993. Summer Students in Virtual Reality: A Pilot Study on Educational Applications of VR Technology. In A. Wexelblat (Ed.) *Virtual Reality, Applications and Explorations*. Cambridge, MA: Academic Press Professional.
- BYRNE, C. M. 1996. *Water on Tap: The Use of Virtual Reality as an Educational Tool*. Unpublished Ph.D. thesis, University of Washington, College of Engineering.
- CHAPIN, W. L., LACEY, T. A. AND LEIFER, L. 1994. DesignSpace: A Manual Interaction Environment of Computer Aided Design. In *Proceedings of the ACM's SIGCHI 1994 Conference: CHI'94 Human Factors In Computing Systems*.
- DURLACH, N. I., ALLEN, G., DARKEN, R., GARNETT, R. L., LOOMIS, J., TEMPLEMAN, J., AND VON WIEGAND, T. E. 2000. Virtual Environments and the Enhancement of Spatial Behavior: Towards a Comprehensive Research Agenda. In *PRESENCE -Teleoperators and Virtual Environments*, Vol. 9, No. 6, pp. 593-615.
- GITTLER, G. AND GLUCK, J. 1998. Differential Transfer of Learning: Effects of Instruction in Descriptive Geometry on Spatial Test Performance. *Journal of Geometry and Graphics*, Volume 2, No. 1, 71-84, 1998.
- JACKIW, N. 1995. *The Geometer's Sketchpad Version 3*. Key Curriculum Press, Berkeley.
- KAUFMANN, H., SCHMALSTIEG, D. AND WAGNER, M. 2000. Construct3D: A Virtual Reality Application for Mathematics and Geometry Education. *Education and Information Technologies* 5:4, pp. 263-276.

- KIYOKAWA, K., TAKEMURA, H., AND YOKOYA, N. 2000. SeamlessDesign for 3D Object Creation, *IEEE MultiMedia magazine*, 7 (1), pp. 22-33.
- LABORDE, J. M. AND BELLEMAIN, F. 1998. *Cabri-Geometry II*. Texas Instruments. Copyright Texas Instruments and Universite Joseph Fourier, CNRS. URL: <http://www-cabri.imag.fr/index-e.html>
- LARSON, P., RIZZO, A.A., BUCKWALTER, J.G., VAN ROOYEN, A., KRATZ, K., NEUMANN, U., KESSELMAN, C., THIEBAUX, M. AND VAN DER ZAAG, C. 1999. Gender Issues in the Use of Virtual Environments. In *CyberPsychology and Behavior*, 2(2).
- MAIER, P.H. 1994. *Raumliches Vorstellungsvermogen*. Peter Lang GmbH, Europaische Hochschulschriften: Reihe 6, Bd. 493, Frankfurt am Main.
- MANTOVANI, F. 2001. VR Learning: Potential and Challenges for the Use of 3D Environments in Education and Training. In *Towards CyberPsychology: Mind, Cognitions and Society in the Internet Age*, Giuseppe Riva & Carlo Galimberti (Eds.), Amsterdam, IOS Press.
- MECHLING, R. 2000. Euklid Dynageo. <http://www.dynageo.com>
- MINE, M. 1996. *Working in a Virtual World: Interaction Techniques Used in the Chapel Hill Immersive Modeling Program*. UNC Chapel Hill Computer Science Technical Report TR96-029.
- MINE, M. 1997. ISAAC: A Meta-CAD System for Virtual Environments. *Computer-Aided Design*, 29(8).
- OPENCASCADE 4.0. 2001. Open-Source Toolkit for 3D modeling. URL: <http://www.opencascade.com>
- OSBERG, K. 1997. *Spatial Cognition in the Virtual Environment*, Technical R-97-18. Seattle: Human Interface Technology Lab.
- REITMAYR, G. AND SCHMALSTIEG, D. 2001. Mobile Collaborative Augmented Reality. In *Proceedings of the 2nd ACM/IEEE International Symposium on Augmented Reality (ISAR'01)*, New York NY.
- RICHTER-GEBERT, J. AND KORTENKAMP, U. H. 1999. *The Interactive Geometry Software Cinderella: Version 1.2 (Interactive Geometry on Computers)*. URL: <http://www.cinderella.de/>
- RIZZO, A.A., BUCHWALTER, J.G., NEUMANN, U., KESSELMAN, C., THIEBAUX, M., LARSON, P., AND VAN ROOYEN, A. 1998. The Virtual Reality Mental Rotation Spatial Skills Project. In *CyberPsychology and Behavior*, 1(2), pp. 113-120.
- SCHMALSTIEG, D., FUHRMANN, A., HESINA, G., SZALAVARI, Z., ENCARNOCAS, M., GERVAUTZ, M. AND PURGATHOFER, W. 2002. The Studierstube Augmented Reality Project. In *PRESENCE-Teleoperators and Virtual Environments*, MIT Press.
- SZALAVARI, Z. AND GERVAUTZ, M. 1997. The Personal Interaction Panel - A Two-Handed Interface for Augmented Reality, *Computer Graphics Forum*, 16, 3, pp. 335-346.
- TAXEN, G. AND NAEVE, A. 2001. CyberMath: Exploring Open Issues in VR-Based Learning. SIGGRAPH 2001 Educators Program. In *SIGGRAPH 2001 Conference Abstracts and Applications*, 49-51.
- WINN, W. 1997. The Impact of Three-Dimensional Immersive Virtual Environments on Modern Pedagogy. *HITL Technical Report R-97-15*. Discussion paper for NSF Workshop. Human Interface Technology Laboratory, University of Washington.



# Assignment: Scene Graphs in Computer Graphics Courses

Dennis J. Bouvier  
Saint Louis University  
djb@acm.org

## Abstract

A number of published papers recommend teaching scene graphs in the introductory computer graphics course [Bouvier 2002; Cunningham 1999; Hitchner and Sowizral 1999; Wolfe 1999]. However, little has been published concerning how to effectively use scene graphs in the introductory computer graphics course. This paper summarizes possible scene graphs exercises and teaching experience of the author.

## 1 Introduction

Scene graphs are data structures representing 3D geometry and other scene properties (e.g., lights and viewer position) in a hierarchical fashion. Scene graphs have been around for years in both academic and commercial settings [Brown 1995]. For several years, scene graphs were primarily used in Open Inventor [Ames et al 1997; Foley et al 1996; Wolfe 1999]; however, the recent popularity of the Java 3D API has attracted more attention to the use of scene graphs for both programmers [Baker 2002; OpenRM 2002; Sun 2002] and educators [Bouvier 2001; Bouvier 2002; Brown 1985; Cunningham 1999; Cunningham and Bailey 2001; Hitchner and Sowizral 1999; Wolfe 1999]. Most recently, open source projects have developed scene graph APIs as well [OpenRM 2002; Open Scene Graph 2002; OpenSG 2002].

Several educators advocate the use of scene graphs for a variety of pedagogic reasons [Brown 1985; Bresenham 1999; Cunningham 1999; Cunningham and Bailey 2001; Hitchner and Sowizral 1999; Wolfe 1999]. However, even within the publications advocating the use of scene graphs in the classroom, little advice has been given on how to use them in the classroom. This paper addresses this void.

## 2 Why Use Scene Graphs

Some may argue, 'one must know about scene graphs to get a job in graphics.' Others might argue that scene graphs are the way of the future for computer graphics. Neither of these reasons warrants the inclusion of scene graphs in graphics courses. However, there are good reasons as put forward by Cunningham [1999; 2000] and Wolfe [1999] as well as others [Cunningham and Bailey 2001; Hitchner and Sowizral 1999].

Scene graphs are useful in a computer graphics course as an abstraction of the 3D graphics rendering system. The educational power of the scene graph lies in the ease with which a complex scene can be represented and the corresponding ability to create and comprehend these complex scenes.

Scene graph concepts are useful in a course even when a scene graph API is not used in programming projects. Scene graphs are useful as design tools and documentation when the concepts are taught and used in the course.

## 3 What is a Scene Graph

A detailed technical explanation of scene graphs would go on for (possibly hundreds of) pages [Ames et al 1997; Cunningham and Bailey 2001; OpenRM 2002; Wernecke 1994]. Succinctly stated,

a scene graph is a data structure and/or graphical representation of a hierarchical 3D scene. The data structure is typically a directed acyclic graph (DAG) in which the nodes of the DAG represent geometric transformations, geometry, lights, the viewer, and possibly other features of the 3D scene.

A particular scene graph implementation or product may be a file format (e.g., VRML [Ames et al 1997]), a programming API (e.g., Java 3D [Sowizral et al 1995]), or both (e.g., Open Inventor [Open Inventor 2002]). Scene graph files require a viewing or display program while an API is used to create an independent application. Section 4.3 presents more on this topic.

### 3.1 A Simple Scene Graph Representation

For the purposes of illustrating some of the assignments suggested in this paper, a simple scene graph representation is presented here. For a more complete explanation, see the cited references at the conclusion of this paper.

A scene graph is composed of at least three kinds of nodes: the root, transform, and visual object nodes. Figure 1 shows one possible set of graphical representations of the three basic scene graph nodes. The organization of these nodes in a DAG represents a scene graph.

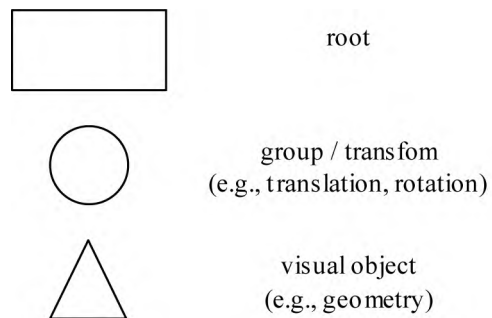


Figure 1 Simple Scene Graph Node Representations

The root node establishes the point of reference for the other nodes and can be considered the center of the 3D scene. There is only one root node in a scene graph and all other nodes in the scene graph are referred to directly, or indirectly, from the root node.

The group node can simply be a collection of references to child objects or transformations of its child object(s) (alternatively viewed as scaling, rotating or translating the reference space for the child objects).

The visual object node is geometry, lights, or other visual content. For the purposes of this discussion, the visual object will always be geometry. One detail being ignored is how material properties (i.e., reflectance coefficients and surface normal vectors) are specified in a scene graph.

A scene graph is typically organized into two major branches: the content branch and the view branch. Having the viewer information in a separate branch of the graph emanating from the

root of the scene graph allows the view position to move (or remain stationary) independently of the content.

Also missing from this discussion of scene graphs are dynamic elements (i.e., nodes for interactivity and animation) and other details, which vary by implementation. The missing details are not necessary to this discussion. The interested reader is referred to the appropriate documentation [Ames et al 1997; OpenRM 2002; OpenSG 2002; Sowizral et al 1995; Wernecke 1994].

### 3.2 A Simple Scene Graph Example

Consider a 3D scene consisting of a robot arm and a cube. The robot of this example consists of a base, a vertical shaft, a horizontal arm, and two 'fingers.' Figure 2 shows one such scene. The x-, y-, and z-axis orientation is shown for reference in this discussion, but is not part of the application.

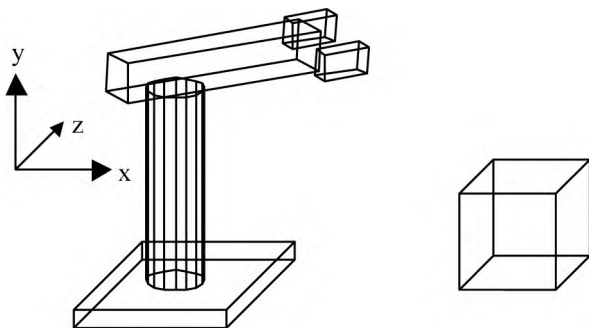


Figure 2: Rendering of Robot Arm with Cube

Figure 3 shows one possible scene graph for the scene of Figure 2. The scene graph is composed of a root node, eight group nodes, and six visual object nodes. The visual object nodes in this scene graph represent geometry of the scene.

As explained above, group nodes may transform their children. Consequently, the composition of the scene graph indicates which visual objects move independently and which move in a coordinated (i.e., hierarchical) fashion. For example, the cube may move independently of the robot (and vice-versa) since the cube visual object is a child of group 2 and none of the robot is. On the other hand, group 3 can move the entire robot, while group 5 can move just the upper parts of the robot.

Of course, other scene graph compositions are possible. However, not all scene graph compositions will allow the appropriate hierarchical behaviors. To design a proper scene graph composition requires having the knowledge (as specified, or through application understanding) of the hierarchical relationships of the scene's contents.

## 4 Scene Graphs in Graphics Courses

When adding new material to a course the instructor is faced with the issues of how to get the students to learn the new material. Adding a new topic to an existing course should be more than adding a lecture or some new reading material. Homework assignments, programming projects, and test questions that correspond to the new material is critical for learning.

Adding scene graphs to a computer graphics course is no different. The instructor considering adding scene graph material to a computer graphics course needs educational activities and testing. Obviously, a student can be required to use some scene graph API to create a graphical program, but what else is there?

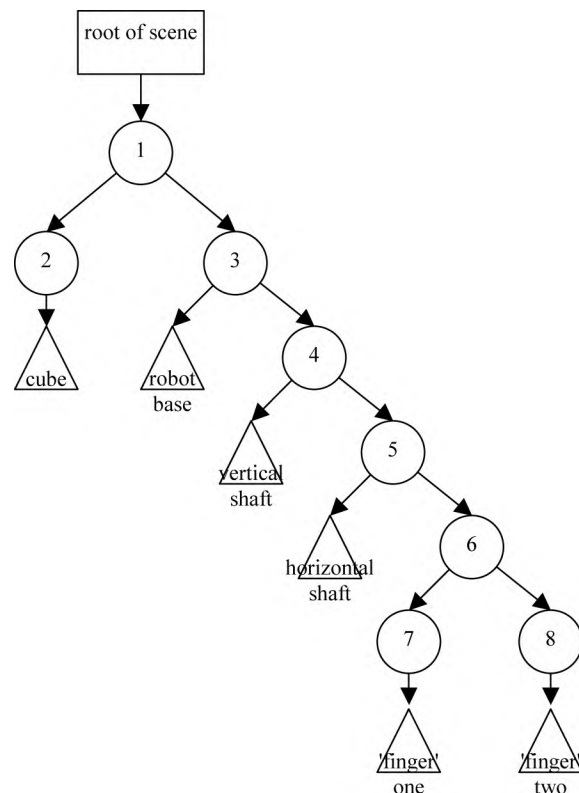


Figure 3: Scene Graph of Robot Arm with Cube

### 4.1 Possible Student Assignments

Below are four possible student assignments involving scene graph knowledge. A course could use all four, but typically assignments A, B, and C would be used for undergraduate courses while assignment D is probably most suited for graduate level instruction.

#### Assignment A: Create a Scene Graph

The first type of assignment is similar to the simple example given in Section 3.2 above. The student is given the description of a scene and asked to design a scene graph. Again, to design a proper scene graph composition requires having the knowledge of the hierarchical relationships of the scene's contents. Consequently, these details should not be absent from the scene description.

A student can undertake this type of assignment very quickly after having been exposed to the basic scene graph information. The student scene graphs can be hand drawn or involve computer tools.

This assignment serves to reinforce the basic concepts and issues involved in hierarchical compositions and (hopefully) increases understanding. More challenging assignments of this type include more advanced scene graph issues such as viewing, lighting, scoping, behaviors and interaction.

#### Assignment B: Scene Graph 'Understanding'

In this assignment a student produces a procedural representation for a given scene graph. This type of assignment can be undertaken only when the student knows about both scene graphs and some low-level implementation mechanism (e.g., OpenGL).

**Table 1 A comparison of Scene Graph features**

API/Language feature	Open Inventor	Java 3D	OpenSG	Open Scene Graph	Performer	VRML
API or Format	API	API	API	API	API	Format
API Host Language	C++	Java	C++	C++	C/C++	N/A
Documentation	Spec., tutorials, books	Spec., tutorials, books	Spec., tutorials	Spec.	Spec., programmer's guide	Spec., tutorials, books
Version (as of 3/1/02)	2.1	1.3	1.0	0.8	2.5	VRML97
Source	Open	Closed	Open	Open	Closed/Open	N/A
Low-level	OpenGL	OpenGL or D3D	OpenGL	OpenGL	OpenGL	None
Organization	SGI	SUN	OpenSG	OpenSceneGraph	SGI	Web3D
OS Platforms	Unix	Windows & Unix	Windows & Unix	Unix	IRIX/Linux	Many

If a student knows about a low-level mechanism that features a transformation stack, then the translation is made from the scene graph DAG structure to the order of appropriate transformation operations with the appropriate stack operations. Completing this assignment gives students an appreciation of stack operations. Starting with the scene graph concept is an easy way to show students a meaningful use for the transformation stack.

In completing this type of assignment the student is forced to think about how the matrix stack implementation is useful in interactive or animated applications. Variations on such assignments can include optimization or other advanced issues.

### Assignment C: Using a Scene Graph Language

A student produces a graphical application using a scene graph API or language. The level of difficulty ranges from fairly easy when implementing simple scenes using VRML to fairly complex when doing complex tasks using a scene graph API. The student may be given a specific scene to create or be left to design one (usually under some guidance or constraints).

The project should require the representation of some hierarchical object (e.g., the robot of the previous section) and animation of the hierarchical object. The power of the hierarchical representation may be best illustrated if the same, or similar programming project is also done without a scene graph.

### Assignment D: Implement a Scene Graph System

The implementation of a scene graph system could be undertaken as a semester long project, or parts of it could be implemented as a single assignment. Open source scene graph APIs are of particular interest for this type of assignment (see section 4.3) as they allow implementation of advanced features without having to build an entire infrastructure.

A project of this type can require extensive knowledge of scene graph operations, a low-level graphics API (e.g., OpenGL), and significant programming experience. Consequently, this type of assignment would typically be reserved for graduate students.

## 4.2 Assessment

Assignments A and B lend themselves to in-class test questions. The assessment questions can be similar to the assignments taking

care not to make them excessive in size. For example, give the description of a small application, such as the robot example, and ask the student for a scene graph representation of the scene. Or, give the student a graphical representation of a scene graph and ask the student how it can be implemented, or how elements of the scene graph related to each other.

Proficiency in the tasks of Assignment C can be tested in small programming related test questions including completing some code, modifying some code, questions regarding understanding of some code example, or API detail questions.

Due to the nature of Assignment D, it does not lend itself to assessment as Assignments A and B. However, interesting questions can be posed by asking how some advanced feature might be implemented or what obstacles exist to its implementation. An open question of this sort should be graded with respect to the time allotted for answering.

## 4.3 Selecting a Scene Graph Implementation

A course can utilize scene graph concepts without requiring the use of a scene graph API or language. However, if programming assignments are to be made (i.e., assignments C or possibly D), then a scene graph API or language will probably be selected for class work.

Making a complete comparison of scene graph APIs and languages could involve an extensive exploration of features and performance. However, such a comparison is probably not necessary for making the selection of a scene graph for educational purposes.

There are several factors to consider in selecting a scene graph API that have little to do with the design of the API (e.g., platform, host language, and documentation). Table 1 summarizes the features most likely to be relevant in making the choice.

As previously mentioned, a scene graph implementation is a file format, a programming API, or both. This distinction determines if a particular scene graph definition is suitable for particular educational experiences. Creating a scene in VRML is not the same as creating an application in Java 3D.

The comparison attributes include those distinguishing the execution environments for each scene graph language may execute in which includes the operating system platform, low-level support requirements, and API host language.

The version number can be an indicator of stability or of the breadth of features present. Approach products with version numbers less than 1.0 with care.

Related to the maturity of a product is the level of available documentation. Older products tend to have more resources, such as tutorials and books, while newer products may have little more than the specification document to work from. The level of available documentation can make a difference in student success (or failures).

Whether or not the source code for a scene graph API is freely available can impact the possibilities of some programming assignments. As mentioned in Assignment D, being able to add advanced features (e.g., optimizations) to an existing scene graph API can be a reasonable assignment for upper level students.

Table 1 is not a comprehensive list of scene graph APIs and languages nor a complete comparison of scene graph languages listed. Table 1 is intended to show some scene graph options and decision criteria. Some scene graph languages missing from this list include OpenRM [OpenRM 2002], SSG [Baker 2002], and X3D [Web3D 2002].

## 5 Conclusion

An informal search of web resources found a number of universities making scene graph programming assignments for computer graphics students. The most common assignment required the student to create a graphics application using Java 3D. Unfortunately, the assignments found represent a small fraction of computer graphics courses. Further, simply creating a scene graph does not reveal the power of this representation.

The student assignments and other information, including the comparative list of scene graph choices, presented in this paper should enable additional instructors to use scene graph material in their courses.

The author has exercised assignments A, B, and C in undergraduate courses over the last two years with some success. Students utilizing scene graphs representations for 3D scenes are able to produce more complex animations with less work. The author is anxious to have graduate students so that assignment D can be tried.

## References

- AMES, A., NADEAU, D. AND MORELAND, J. 1997. *VRML 2.0 Sourcebook*, Wiley.
- BAKER S. 2002. SSG: A Simple Scene Graph API for OpenGL, <http://www.woodsoup.org/projs/plib/ssg/>.
- BOUVIER, D. 2001. From Pixels to Scene Graphs: The Evolution of Computer Graphics Courses, *SIGGRAPH 2001 Conference Abstracts and Applications*, August 2001
- BOUVIER, D. 2002. From Pixels to Scene Graphs in Introductory Computer Graphics Courses, to appear *Computers and Graphics*.
- BROWN, M. 1985. *Understanding PHIGS: The Hierarchical Computer Graphics Standard*, Template Software Division of Megatek Corporation.
- BRESENHAM, J. 1999. Teaching the Graphics Processing Pipeline: Cosmetic and Geometric Attribute Implications, *Proceedings of Graphics and Visualization Education 1999*, July 1999.
- CUNNINGHAM, S. 1999. Re-Inventing the Introductory Computer Graphics Course: Providing Tools for a Wider Audience, *Proceedings of Graphics and Visualization Education 1999*, July 1999.
- CUNNINGHAM, S. 2000. Powers of 10: The Case for Changing the First Course in Computer Graphics, *The Proceedings of the Thirty-first Technical Symposium on Computer Science Education*, Austin, TX, March 2000, 46-49.
- CUNNINGHAM, S. AND BAILEY, M. 2001. Lessons from Scene Graphs: Using Scene Graphs to Teach Hierarchical Modeling, *Computers & Graphics* 2001, number 4.
- FOLEY, J. ET AL 1996. *Computer Graphics: Principles and Practice in C, 2/e*, Addison Wesley.
- HITCHNER, L. AND SOWIZRAL, H. 1999. Adapting Computer Graphics Curricula to Changes in Graphics, *Proceedings of Graphics and Visualization Education 1999*, July 1999.
- OPEN INVENTOR 2002. <http://oss.sgi.com/projects/inventor/>.
- OPENRM 2002. OpenRM Project Overview <http://openrm.sourceforge.net/overview.shtml>.
- OPEN SCENE GRAPH 2002. <http://openscenegraph.org>.
- OPENSIG 2002. <http://opensig.org>.
- PERFORMER 2002. *SGI-OpenGL Performer Overview*, <http://www.sgi.com/software/performer/>.
- SOWIZRAL, H., RUSHFORTH, K. AND DEERING, M. 1995. *The Java 3D™ 3D API Specification*, Addison-Wesley.
- SUN 2002. *Java 3D API home page*, <http://java.sun.com/products/java-media/3D>.
- WEB3D CONSORTIUM 2002. *X3D home page*, <http://www.web3d.org/x3d/>.
- WERNECKE, J. 1994. *The Inventor Mentor*, Addison-Wesley, 1994
- WOLFE, R. 1999. Bringing the Introductory Computer Graphics Course into the 21st Century, *Proceedings of Graphics and Visualization Education 1999*, July 1999.

# Building Computer Graphics Education in Developing African Countries

Sampson D. Asare  
University of Botswana  
asaresd@mopipi.ub.bw

Petros M. Mashwama  
University of Swaziland  
petros@science.uniswa.sz

Steve Cunningham  
California State University Stanislaus  
rsc@cs.csustan.edu

## Introduction

Computer graphics is becoming a tool for communication, change, and development throughout the world, and in order for this tool to be effective, there must be an educational base for its use. In southern Africa the development of computer graphics is underway, but it has not made the impact it could have. We believe that one of the main reasons for this is that there has not been as much educational development as is needed in the region. This note will describe some of the challenges that are faced by educational institutions in southern Africa as they try to develop the computer graphics education in the region, and will discuss some of the ways these institutions are working to meet these challenges.

As is the case in any other part of the world, there is a considerable range of educational facilities and opportunities in the region. There are historically strong universities in South Africa, expanding universities in developing countries such as Botswana, Namibia, and Swaziland, and polytechnic institutes throughout the region. While some of the challenges we discuss may apply more to one or to another kind of institution, or even to the special case of an individual institution, they all have an impact at some level.

## 1. Challenges

Developing computer graphics education involves updating and creating courses or course content and, perhaps, removing course content in other areas because curriculum is often a zero-sum game: in order to add something, something else must often be removed. There is always a challenge in making this kind of change to a curriculum, but there are some additional problems when we introduce or expand a topic such as computer graphics into a computer science curriculum in a developing country. Some of these problems are shared by computer science programs across the developed world as well, but are especially difficult in the developing world. We have identified three major kinds of problems, each of which has particular aspects that we will describe.

### 1.1. Student preparation

For many reasons, from limitations in the secondary education that has been available to students to a lack of emphasis on the subject, students come to their university studies with a very limited mathematical background. This is, of course, a problem that is not limited to developing nations, but here it goes beyond limitations in student interest to a limitation in student opportunity. With this lack of background, students have difficulty with mathematical topics and seem to lose interest when such topics come up in computer graphics, as they necessarily must.

A similar problem comes up with programming skill. Again, the limitations in mathematical and other analytic studies at the primary and secondary level leaves students with limited

background for programming, which leads to limited programming skill and experience. Computer graphics in the computer science context is quite programming-intensive, and without programming skill it is difficult to focus on the geometric and spatial skills, especially the 3D thinking skills, that students must develop.

With these two preparation problems, there is a significant problem with student motivation. It is difficult to motivate students to study computer graphics on its own because without seeing what can be accomplished with computer graphics, especially what can be done that is applicable to their own lives, students do not develop the passion for the subject that is needed to sustain their learning of sometimes difficult concepts and processes.

### 1.2. Isolation

It is a fact of life in southern Africa, as it is in many other developing regions of the world, that the professionals in the region often suffer from an isolation from the progress and new approaches to problems that may be found in more developed regions. This is part of the classical “digital divide” problem that has such a strong impact around the world. When one works in a field with the rapid changes and constantly-growing ideas of a topic such as computer graphics, this is a critical problem with many aspects.

One problem caused by isolation is that someone working in computer graphics will often have very few people to talk with about the field. In Botswana, for example, there is a group of people at Botswana TV who work with computer graphics, and these are the primary contact for anyone teaching computer graphics in the country. These are good people, but they do not have experience in scientific or engineering aspects of computer graphics, and they do not approach the subject from a computer science perspective, so they provide a valuable but limited set of contacts for education.

New faculty and new ideas are the lifeblood of a growing educational area, but this isolation makes it difficult to have either. This has two very important impacts on computer graphics education. First, without an active community with connections to the developments across the world, it is difficult to create a computer graphics curriculum that is up to date. Courses tend to stagnate and it is difficult to integrate new techniques and technologies into instruction. And second, an isolated community can find it very difficult to attract people to teach in the field and can have difficulty retaining anyone who has made the effort to develop skills in the subject.

### 1.3. Limited support

In order to develop any educational emphasis, there must be a great deal of support for it from a number of sources. There must be a national understanding that it is an important area

that serves national goals, there must be an understanding in the educational institution that the area should have the resources necessary to become a strong programme in the institution, and there should be a public understanding that the public will benefit from the contributions of those who study in the field.

Computer science has recently begun to have this kind of support in developing countries, and university programmes are beginning to get improved facilities and faculties. At the University of Swaziland, for example, laboratory facilities are being rebuilt and new PCs installed, while at the University of Botswana there are good student PC laboratories but these do not have Internet access because it is very expensive and the bandwidth is very limited. Both universities have solid and expanding faculties in computer science.

This support for computer science does not necessarily extend to computer graphics, however. The support for computer graphics in developed countries has been built over the last 15 years from the days when scientific visualization and other graphics applications were seen as ways to make nice pictures but as only marginal to the success of other fields, to the present day when almost all enterprises of any importance include computer graphics as a key part of their development. The key to this changed perception is the 1987 report on visualization in scientific computing [McCormick 1987].

But in developing countries, this support for computer graphics has not yet been achieved. More traditional economies and the limited impact of technologically-based activities make the promise of computer graphics more theoretical than real. And without this support, it is difficult to make the case for increased financial and human resources to develop instruction in computer graphics. Most institutions except the largest will have only one (if any) faculty member who teaches computer graphics, and many institutions are only now starting their first computer science course in computer graphics. Even within education, there is a very limited use of computer graphics in the instruction in other fields or as a component of distance learning.

## 2. Approaches to improvements

While the challenges are significant and in many ways are daunting, there are developments in southern Africa that have the promise of making a very large difference in the importance of computer graphics in the region. We believe that these developments will allow us to see a growth of computer graphics education and application within the next few years that will make computer graphics a major force in the development of the region.

The most important development is the creation of the African Association for Computer Graphics, or Afrigraph [Afrigraph 2001]. This association was created in 2001 with the support of ACM SIGGRAPH and of Eurographics [Chalmers 2001], and its goal is to serve computer graphics professionals in Africa by providing a community of persons in the field, a conference where persons from Africa and persons from other parts of the world can meet to exchange ideas, and a contact for two-way sharing of information between Africa and the rest of the world through relationships between professional associations.

The immediate effect of Afrigraph and of its first conference (at Camps Bay, near Cape Town, South Africa) was to create a great

deal of interaction among people from southern Africa and other countries, to share information on the computer graphics activities in the region, and to discuss potential collaborations in research and education. [Brown 2001; Chalmers 2001] In particular, a panel on education, organized by the authors of this paper, generated a great deal of discussion that was the primary source of the ideas we present here. By sharing their experience, computer graphics instructors were able to see that they had common problems and to share ideas on how they might be overcome. In addition, computer graphics professionals within Africa were able to meet new people from other countries and to share their experience—because there is a significant amount of computer graphics activity within the region—with others.

In addition to sharing information at the Afrigraph conference, the relationship between Afrigraph and both ACM SIGGRAPH and Eurographics will help keep open the communication channels between Africa and the rest of the world. By sharing publications and materials from the different associations' conferences, members of all these associations will have a broader understanding of the developments around the world. And by sharing personal connections, Afrigraph can help with the academic exchanges and with the visibility of computer graphics activities that can help develop the skills of current faculty and attract new faculty to the region.

For education, the discussions at the Afrigraph conference provided a number of ideas. It was seen that there could be ways to introduce students to computer graphics without needing so much mathematics that their enthusiasm was lost. Including such innovative ideas as learning drawing skills, students can learn to see spatial relationships. By focusing on some of the practical aspects of computer graphics, students can see the value of the subject and will develop a passion for the subject that can carry them through later studies that are more mathematical. By seeing the value of computer graphics as a way to understand other topics, students can see that the subject has important contributions to other fields. It was generally agreed that computer graphics studies should begin as early as possible, but could not start before the student had developed programming skills, and that there should probably be at least two undergraduate courses, one more practical and based on a graphics API, and the next more theoretical and mathematical. For one possible approach to this combination of courses, see [Cunningham 2000]. There is also an interest in developing educational opportunities in multimedia to take advantage of the ability of computer graphics to support communication and to assist in the teaching of other subjects.

In order to develop the resources needed for education, computer graphics instructors will need to reach out to others to show the value of our subject. We can take advantage of any joint programs in our universities to share our knowledge with other academic programs or other institutions, and in particular we can help faculty from the sciences, engineering, or other areas to address problems in their fields with our tools. When other fields start using computer graphics in their teaching or research, it will increase the motivation of students to study computer graphics and will give us a better case for increasing our financial and staffing support.

Beyond this, though, we have the opportunity to help make computer graphics a key part of the development of our countries and our region. It seems particularly important to use computer graphics to help address the particular problems that

will make a difference to the public and to the government, because this will ensure that our universities will understand the importance of computer graphics education to our national interest. This will need the collaboration of computer graphics with other academic fields as well as with non-academic activities, and will stretch the already-thin computer graphics resources, but will pay large dividends over time.

### 3. Conclusions

There are rich opportunities for computer graphics to contribute to the development of southern Africa. There is untapped talent among students and there are dedicated faculty who care deeply about helping students achieve success in the field. The significant challenges of student preparation, keeping education current with developments in the field, and developing support for the field in the region can be met, and the African Association for Computer Graphics will play a key role in helping computer graphics professionals in the region keep up with each other and with the activities of the rest of the world. When this is combined with focused efforts to have computer graphics contribute visibly to addressing the problems of the region, we are very optimistic about the future of computer graphics in southern Africa and in Africa as a whole.

### 4. References

- AFRIGRAPH. 2001. Afrigraph organization and conference pages, retrieved March 11, 2002 from <http://www.afrigraph.org/>
- BROWN, J. 2001. Report on the Afrigraph conference, retrieved March 11, 2002 from <http://www.siggraph.org/conferences/REPORTS/Afrigraph2001.html>
- CHALMERS, A. 2001. Report of the ACM SIGGRAPH visit, retrieved March 11, 2002 from <http://www.siggraph.org/project-grants/SA/sigsareport.html>
- CUNNINGHAM, S. 2000. Powers of 10: The Case for Changing the First Course in Computer Graphics, *Proceedings of the SIGCSE 2000 conference*, Austin, TX, March 2000, pp. 293-296
- MCCORMICK, B., DEFANTI, T, AND BROWN, M, eds. 1987. *Visualization in Scientific Computing*, *Computer Graphics* 21(6), November 1987

# Macromedia Flash in Physics Education

## ASPIRE's Interactive Online Labs and Lessons

Julie Callahan  
University of Utah  
Julie@cosmic.utah.edu

Charles C. H. Jui  
University of Utah

### Abstract

Flash provides the ASPIRE team a means to create interactive virtual learning environments, where teachers and students can explore places and things previously limited by the classroom or simply by time constraints. Flash can become an immersive world for online science education. Producing curriculum support material in Flash provides many benefits for teachers and students.

Teachers are provided with an affordable tool, which aligns to local and national curriculum standards. This material provides a rich classroom experience at little cost to the educator, while also satisfying the curriculum goals. Students can participate in an engaging learning experience, where science becomes an experience beyond a textbook lesson or a classroom lab. Situations are presented that would be either difficult to reproduce due to cost or scale. For example, students can investigate the causes of tides in the ASPIRE Flash activity "Gravity and Tides."

Flash, the software available for authoring these activities, can produce a mathematically correct, physically based model that students can observe and investigate. These lessons provide a high-quality experience; meanwhile, the cost and speed of production of these activities continue to decrease.

### 1 Flash and ASPIRE Lessons

Flash technology has opened the Web to a vast number of interactive sites that have been designed to sell, promote, dazzle and impress. Indeed, designing a multimedia site has now become an art form exclusive to the Internet. When the team at the Astrophysics Science Project Integrating Research and Education, or ASPIRE, saw Flash in action, they saw an opportunity to create something different and new. Flash can become an immersive world for online science education. Initial production of ASPIRE Flash lessons created immediately observable benefits, including greatly reduced production time and cost. Additionally, Flash productions increased the level of interactivity and production quality.

Flash provided the ASPIRE team a means to create interactive virtual learning environments, where teachers and students can explore places and things previously limited by the classroom or simply by time constraints. As an astrophysics outreach program, the ASPIRE team determined that the lessons best suited for online visualization were based on Earth Systems curriculum dealing with astronomical subjects. Upon completion, this first set of lessons would create a 'virtual planetarium.' This production would follow a modified instructional development model, customized to use local and national curriculum standards. [Gentry 1994; USOE 1995]

"Gravity and Tides" is the first set of Flash lessons that the ASPIRE team completed as part of the virtual planetarium. The immersive quality of the lessons exceeded the expectations of the ASPIRE team and work immediately progressed to other subjects. Flash is able to integrate basic physical formulas to produce simulations. These simulations result in an accurate model of

gravity and tides in our local solar system. Because the simulation is mathematically based, the student can interactively manipulate the model to explore variations of the model. The virtual lab provides a rich hands-on experience that a teacher would be challenged to reproduce in the classroom. These lessons are unique in that they have integrated curriculum and activities. They are also free of charge for anyone who wishes to use them.

### 2 ASPIRE

ASPIRE is the educational outreach program for the High Resolution Fly's Eye Cosmic Ray Research Group (HiRes) at the University of Utah. The HiRes research group has operated at the University of Utah since 1991, and is funded by the National Science Foundation. In 1997, the NSF charged HiRes scientists with forming an educational outreach program. A committee of scientists, educators and administrators held a summit to determine how HiRes could best serve K-12 educators and administrators.

At the end of the meetings, it was determined that the greatest need for science education outreach was at the middle school level. According to studies by Project 2061, a national science education reform panel, textbooks for science teachers and students at the middle school level have been found to be unsatisfactory. [Roseman et. al. 2001] Recreating textbook situations was neither the answer to this complex issue, nor is a visit practical or possible by research scientists to an area as large as the Utah. A consensus was reached to provide the 7th, 8th and 9th grade teachers with online activities, so that they can benefit from the outreach program, regardless of proximity to the University of Utah.

A team was assembled with teachers, scientists, programmers and artists to produce web-based activities. The first ASPIRE lessons were produced as applets written in Java. One unique feature to these lessons that is still difficult to find on the web today is that the curriculum is integrated into a web page that serves both as lesson and lab. Instead of having a lecture day and then a lab day, the student experiences both on the Web. These lessons were well received and continue to be used in schools in Utah as well as nationally and internationally.

The benefits of using Java were immediately obvious. Students could virtually interact with the computer in a simulation that could not be easily or affordably reproduced in a classroom. Computers have an appeal that engages students, who otherwise may not participate in the classroom. Teachers could spend valuable classroom time answering the students' questions about the lessons instead of lecturing. The lessons developed were designed to adhere to national science education reforms, so that any teacher could feel comfortable utilizing these activities and know that the students were exploring concepts and ideas that may be found on an end-of-level test.

While the lessons were successful, Java presented some drawbacks. These included technical difficulties for the users. Java is expensive to program as well. ASPIRE hires students at the University of Utah to provide an enriched work experience. A



considerable amount of time and budget was devoted to learning Java. At that time, the benefits of the science lessons and labs far outweighed the drawbacks.

### 3 FLASH

In October of 2000, Macromedia released Flash 5. The previous versions of Flash were used by many designers to create exciting web pages and interactive animations, but had not been considered by ASPIRE for more intensive programming functionality. With Flash 5, however, the new ActionScript language allowed the additional level of flexibility needed to create interactive science labs. The most recent browsers were being released with built-in support for the Flash viewer as well, which eliminated some of the previous technical concerns. ASPIRE designed a production model for a preliminary activity, and the Tides and Gravity Simulator was developed.

#### 3.1 Creating a Realistic, Physically-based Tide Simulator

The first completed activity was the Tides and Gravity lessons. This unit is actually a set of 4 separate lessons and labs written by 8<sup>th</sup> grade science teacher, Gina Ward. Production started with a storyboard session to determine how the activity should progress, and what level of interactivity the students could have with the lab portion of the lesson. Tides are the result of planets and satellites in motion and their mutual gravity in our solar system. They can be mathematically explained using Newton's Universal Law of Gravitation:

$$F = \frac{GMm}{r^2}$$

How could the ASPIRE team turn this mathematical formula into an exciting interactive lab for 9th grade students? Students in a land-locked place like Utah rarely get to observe tides and students all over would have a hard time seeing what tides would look like from outer space.

The lab would allow students to 'see' what tides look like, if they had a birds-eye view of our solar system. The tide movie was created in Flash using the new capabilities that are now available through ActionScript. The concept of the movie is simple. There are three rotating objects: the Sun, Moon, and Earth; the Earth is surrounded by molecules of water. Due to the complex math between each object on the screen animating and updating each frame, the number of water particles is limited.

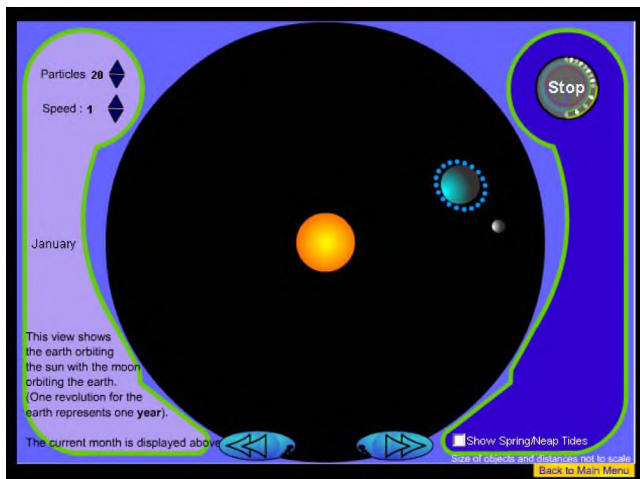


Figure 1 ASPIRE Tide Simulator

This interactive lab, combined with the curriculum lesson written by Gina Ward, creates a rich classroom experience. The students can explore and observe the high tides and low tides that happen as a result of lunar cycles. They can examine and explore why there is a bulge not only facing the moon, but also on the opposite side. As a guided inquiry-based activity, another version of this lab was created where a second moon could be added. The students hypothesize what they think will happen when an additional moon is added to the Earth's orbit, and then observe the results in the next simulation. Because the Flash movie is based in the model, adding another body is a simple exercise.

After the tides lesson and lab was created, other activities were designed to complement the "Tides Simulator". Now when students explore "Gravity and Tides", the first activity is an exploration in critical speed. The second activity is actually an exploration in how Newton's Universal Law of Gravitation works, without having to crunch numbers, but rather observe the relationship of the formula relative to the sun, Earth, and moon. The tide simulator is the third activity, followed by a final activity, where the student takes real, historical data from Monterey Bay over the course of a year.

The result of this activity is a rich interactive experience. The level of immersion that this activity offers is greater than ASPIRE had been able to produce prior to working in Flash. The movies can accept text input fields and dynamically interact with the text. With the previous online lessons, students worked from a printed lab book, and converted data taken on the computer to their paper booklet. The dynamic text function allows students to answer questions, hypothesize and receive immediate interaction from the computer. Flash can also write a text file to be printed so homework can be printed out and submitted.

Exploring the lessons becomes more than just a 'read-and-click' activity. One of the criteria in deciding to develop an ASPIRE activity rests on the necessity of the computer. If a teacher can create by other means a classroom experience just as easily and inexpensively as an ASPIRE lesson, then there is little point in producing it. If a lesson can save a teacher time and money, then value is found in creating the activity.

The artwork and environment can help create an environment where the student can read and explore; meanwhile the experience is enriched with the interactivity. Simply reading on the computer presents the same problems as simply reading a textbook. Experiencing a concept or idea engages students, and an engaged student is a learning student.

Other benefits in the production model became immediately obvious. The method of working between programmers and artists is more closely connected. Programmers can set up a file system with placeholder artwork, which the artist can then incorporate. Java programmers had difficult enough time explaining their code to each other, let alone allowing the artist to load and manipulate the files. The communication between programming and artwork becomes closer with Flash. Flash is not only a tool to program, but also an animation and graphics package. This software design creates a production model in which the line between art and technology becomes blurred with use.

The time in producing a Flash lesson has exponentially decreased compared to Java. A single applet in Java took up to a year to program, while the entire "Gravity and Tides" lesson took under a month. Because of this reduction in production time, the cost of each lesson has decreased. To date, four lessons have been published to the web site in support of the virtual planetarium. ASPIRE plans to complete the virtual planetarium by the end of summer 2002, which will more than double the amount of content available on the web site with over 20 individual lessons and labs. This production phase will last a little over a year with a reduced staff of 4 or 5, compared to the body of work that took 3 years to

complete with a larger staff. The overall effectiveness of these activities will be observable after classroom testing and assessment.

#### **4 Conclusion**

Lessons and activities designed for classroom use on the computer can greatly increase the availability of support materials for teachers and students. As newer technologies are developed for authoring multimedia, these types of activities will become increasingly easy and affordable to develop. The lessons that ASPIRE has developed with Flash demonstrate this. ASPIRE has taken steps to provide curriculum appropriate support material that transcends traditional textbook lessons. These methods will support current science education reform efforts while supporting the curriculum standards.

The resulting activities provide many benefits for teachers and students alike. Students can participate in a visually exciting; engaging activity that meshes traditional textbook learning with a hands-on activity. Participation can lead to investigation of subjects previously difficult to visualize or investigate. Teachers have a greater pool of resources to pull from while developing curriculum for the classroom. Because these lessons are digital and online, cost is minimal or free for educators. Curriculum support materials that are aligned with the most current standards will provide valuable resources, both in content and cost, for science educators and students.

#### **References**

GENTRY, C.G. 1994. Introduction to *Instructional Development Process and Technique*. Wadsworth Publishing Co.

ROSEMAN, J. KESIDOU, S. STERN, L. AND CALDWELL, A. 1999. HEAVY BOOKS LIGHT ON LEARNING, AAAS Project 2061 Evaluates Middle Grades Science Textbooks. In *Science Books & Films* 35, 6.

UTAH STATE OFFICE OF EDUCATION 1995. *Earth Systems - 3600 - 1995*. <http://www.uen.org/cgi-bin/websql/lessons/c3.hts?core=3&course num=3600>.

# The Vertex Project: Exploring the Creative Use of Shared 3D Virtual Worlds in the Primary (K-12) Classroom

Fiona Bailey  
Middlesex University  
f.j.bailey@mdx.ac.uk

Magnus Moar  
Middlesex University  
m.moar@mdx.ac.uk

## Abstract

Children invent imaginary worlds and enact scenarios within them on a daily basis as part of their imaginative play. Given the opportunity and the tools, what kind of worlds would children create for themselves within a virtual space, and what kind of learning can emerge within these playful, child-centered spaces?

In VERTEX, young children inhabit an imaginary virtual world that they have designed and created using 3D modeling tools and net-based virtual worlds software. Crossing traditional subject disciplines and involving local and remote collaboration, the project demonstrates children's design and communication abilities above and beyond the expectations of the curriculum.

## 1 Introduction

There is a growing body of research demonstrating that shared virtual worlds can be powerful educational tools with the potential for transforming the learning experience. Recent studies carried out by researchers at the Electronic Visualization Lab at UIC [Roussos et al. 1997; Johnson et al. 1999] using immersive virtual reality have indicated a range of positive effects on learning, as has work underway at MIT's Media Lab [Bers 1999] using desktop virtual environments. They can present engaging interactive environments for learners to explore, can encourage active participation collaboration and exchange, and can be highly effective tools in teaching complex concepts and ideas.

In the UK, however, where many primary schools are still struggling towards the effective integration of the Web into classroom activity, there have as yet been few concrete opportunities for teachers to explore the potential of these technologies. Working within the demands of a highly structured National Curriculum, where the use of computers is often assigned to the subject area of Information Communication Technology (ICT), few teachers have been able to experiment with more creative, cross-curricula methods in relation to new technology.

The aim of VERTEX is therefore to work with teachers and children to investigate the creative teaching and learning possibilities of shared 3D virtual worlds and to develop and disseminate innovative yet practical strategies for their use across the primary (K-12) curriculum.

For the purposes of this study, we have used net-based virtual worlds. Immersive systems are currently large, costly and their use still requires a high level of technical expertise. For many schools operating within limited budgets and with few resources, there are obvious practical difficulties in relation to implementing their use in the classroom at this time. Net-based virtual worlds, on the other hand, are increasingly accessible. They are inexpensive and run easily on a standard PC with a dial-up connection. As most UK primary (elementary) schools now have at least one classroom-based computer, and over 90 percent of schools are now online, this presents us with a realistic opportunity to introduce and develop the use of this technology in a working classroom.

## 2 The VERTEX Project

The starting point for VERTEX is the importance of learning through making. Taking the constructionist idea that learning can occur deeply and effectively through the making of personally meaningful artifacts and projects [Papert 1980] and the belief that imagination and play are fundamental factors in learning [Vygotsky, 1978], VERTEX places children in the role of designers and producers, as well as users, of these media. In doing so, the project aims to monitor and evaluate the learning that arises through children's direct participation in these processes.

Based in three UK primary schools, VERTEX has involved over one hundred 9 to 11 year olds and their teachers in a three-year program of activities toward creating and populating a small network of virtual worlds. These schools are in very different geographic and cultural locations—Soho Parish Primary is at the very heart of London's West End, Oakthorpe Primary located in suburban outer London, and Firth Primary is situated in a small coastal town on the island of Orkney, off the north coast of Scotland. Each school has a diverse social and cultural mix, including children learning English as a second language and children with a variety of special educational needs. They also have varying levels of access, skills and experience in uses of new technologies, amongst the teachers and children.

In order to establish the project effectively within each unique situation, considerable time was spent in the schools during the initial stages. Working closely with teachers, time was spent devising strategies to integrate the project into the class timetable in ways that ensured the full involvement of all children, in all aspects of project work. Working in the context of a classroom of 32 children and having limited access to computers, the best way to achieve this was to develop methods that did not rely solely on computer-based activity.

While the hub of activity centered on the avatar-based virtual world software ActiveWorlds, the project incorporated both digital and non-digital approaches to developing ideas, progressing designs, and 3D object making. Children used new media in combination with traditional activities, ranging from story writing, collage and drawing to the use of software tools including Adobe Photoshop and 3D Studio Max. In each school they worked together as a class, in small groups, and through online teaching sessions towards imagining, planning, designing, and making their own virtual world and a unique set of avatar characters.

From the outset, embedding the project into a working classroom established a sense of ownership within the schools and amongst children and teachers, allowing a clear direction to emerge based on the different character of each school. It also established confidence and an aptitude amongst teachers in handling new tools and managing the project in their classroom.

Working in this cross-disciplinary way, the project specifically aimed to identify the ways in which involvement in the process of virtual world design and construction can develop children's creative expression and imaginative abilities. In addition, we are examining how working in multi-user spaces

could develop their language skills, communication skills and how working with their peers locally and remotely might encourage collaborative learning between young children.

### 3 Developing the Project in School

Work began in January 2000 and following familiarization with the basic functions of the ActiveWorlds interface, children were introduced to their remote partners through the VERTEX world – a secure space separate from the public network. As their skills and confidence grew, children were taught how to construct their own virtual buildings using the duplication and editing tools built into the browser. They also worked together to create structures using lists of downloadable building objects. In the initial sessions, a remote researcher guided children, but as they became proficient, children would meet, explore and experiment together.

Once these skills were acquired, partnerships between schools had become established and children had been instilled with a sense of purpose for creating their environments—the task of developing their creative input could begin. This involved the children collectively devising the overall concept for their virtual world and its various component parts, designing and making their own objects and models that would shape these landscapes, and creating their own avatar characters to inhabit their new worlds.

Initially, children were encouraged to develop their design ideas through activities such as group discussion, creative writing, and role-play. Once a collective plan had been formed, they progressed their ideas through illustrated maps and elaborate drawings. These images formed a working blueprint for the design of the virtual world. During this time, children were also creating their avatar characters. This activity usually began using a ‘hot-seating’ exercise, through which they described their personality and appearance in role. Following this, the children then created designs for their characters in the form of elaborately collaged, movable puppets or as detailed drawings in their workbooks. All these procedures happened as whole class activities directed by the teacher and were linked to various curriculum areas such as literacy and art, as well as ICT, establishing cross-curricular ways of working.

Once the overall picture of each world and its inhabitants emerged, the next stage was to transfer their designs into the computer. In order to do this, they would be using digital photography and 3D modeling tools, namely Adobe Photoshop and 3D Studio Max. Bearing in mind the primary geometric shapes used for building in ActiveWorlds, children mapped out the designs of individual building blocks and objects needed to realize their ideas.

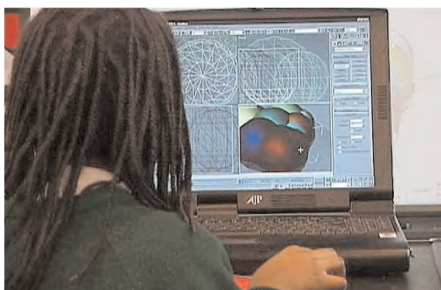


Figure 1: Shanni creating a cloud in 3D Studio Max.

3D Studio Max was introduced to whole class groups using a projector and this first demonstration was used as an opportunity to show children how their puppets were to be transformed into avatars. This involved scanning them into Photoshop, creating a

path and exporting it to 3D Studio Max where it was extruded into a 3D figure. While illustrating this step-by-step procedure, time was spent taking children through the aspects of the program they would need to learn in order to create their building objects. Following these sessions, children were all given the opportunity to play informally with the program. For several weeks following this introduction, children participated in concentrated teaching sessions with the program, as shown in figure 1. To gain the level of competency necessary to create their objects, they needed to learn several basic functions of the program—how to create, combine, and modify simple 3D objects using standard and extended primitives, as well as how to create and map textures onto the objects using the Materials Editor. In order to facilitate this, each child was given the task of creating a tree, made from two or more textured 3D shapes. In preparation for this process, children began creating libraries of images to use as Texture Maps. They did this by participating in ‘Texture Hunts,’ using a digital camera to collect interesting and unusual patterns, shapes and colors from in and around their schools.

As we had anticipated, once working in 3D Studio Max most children understandably struggled with the interface, finding it cumbersome and difficult to manage. They also had difficulty organizing their work spatially. However, it was particularly interesting to observe that the level of excitement associated with creating a ‘real’ virtual object to add to their world created a powerful motivation and impetus for them to persist and ultimately overcome these difficulties. With much repetition, practice, and reinforcement over several weeks, children became more confident and skilled, and began to experiment—making clouds, creatures, furniture, robots, flowers, abstract sculptures, and building blocks to make houses, palaces, and playgrounds. Once completed, their objects were placed on the ActiveWorlds server, where the children could download them and begin the process of constructing their own imaginative shared 3D virtual worlds. Figure 2 shows one such example.



Figure 2: Children dancing in the Astronomic Space Disco.

### 4 Emerging Issues

As we move into the final stages of the project, a number of issues have begun to emerge that we will be investigating more closely. These notably include:

- *Based on work carried out so far, there are strong indications that virtual worlds can be used to great effect as a hub of activity for whole class project work.*

Working in this way can act as a catalyst or an imaginative bridge for children, enabling them to make meaningful links between ICT and other subject areas in the curriculum. An example of this emerged through science topic work while assessing children’s understanding of how animals adapt to their environment. Normally where children would be asked to write about an animal such as the Desert Rat, teacher Judi Hammill

assigned children to design a virtual animal and to describe how this imaginary creature developed its characteristics to suit its unusual habitat. The extensive work produced by the children demonstrated how learning can become more creative when connected to something meaningful and significant to the child's experience.

- *Storytelling and imaginative writing provided a starting point for children to think creatively about their project.*

As work progressed, their activities and encounters in the virtual space began to act as a stimulus for further writing and storytelling in other areas of classroom activity. Teachers working with the project noted that boys in particular, who in general underachieve in literacy at this age (most especially in imaginative writing), showed considerable improvement in their imaginative writing skills when relating work to their virtual world activities. Additionally, teachers have identified the potential benefit for children learning English as a Second Language.

- *Working together in the virtual world demonstrated the potential to significantly develop children's collaborative and communicative skills.*

Through participation in an online project, children learned to skill share and problem solve and worked collectively to develop a sense of joint authorship and ownership. Several children working at one computer has its drawbacks, but there are also advantages as it stimulates conversation about the task. Children by necessity learn to cooperate. In addition, there is potential to form partnerships between significantly different locations. This presents a valuable opportunity for children to explore ideas of difference through sharing stories about each others lives, cultures and experiences

- *There have been strong indicators that children with Special Educational Needs can benefit from working with virtual worlds.*

All children have been highly motivated by working with these technologies, but this has been especially noticeable among children with learning difficulties and children who are otherwise under-motivated or disruptive. This can be attributed to a child-centered approach that involves small group work, but also to the game-like quality of virtual worlds, which engages children playfully and encourages them to keep trying where they may otherwise give up. Teachers have noted a significant increase in the confidence and self-esteem of some children and their achievements in areas such as ICT are developing beyond expectations in relation to the primary curriculum.

## 5 Conclusions

Although there have been many positive outcomes so far, there are still a number of areas in relation to technical implementation that need further investigation and adaptation. Currently, the means by which children build their avatars and objects is through the use of professional authoring tools such as 3D Studio Max and Adobe Photoshop. Whereas there are learning benefits from using such software, there are also several problems:

- There is a considerable time and cost—both for the children in becoming competent with these complex tools and also for the educators in teaching the relevant skills, once they themselves have acquired them.

- The tools are not integrated with the ActiveWorlds browser, which means that children lose engagement with the focal environment that they are building.
- The process currently involves a lengthy delay between children making artifacts, and those objects subsequently appearing in their 3D world. This lack of immediacy can affect the children's enthusiasm and also unnecessarily prolong this aspect of their VERTEX work. This delay is partly a result of the modeling being done outside the browser and also because technical support is required to process the children's models and place them on the ActiveWorlds server.
- The professional authoring tools can be expensive. The ability of most schools to afford them is questionable.

To address these issues, it is proposed that new tools need to be developed. These tools should be fully integrated with the 3D browser, allow direct construction of objects within the 3D world and whilst having simple interfaces, have sufficient functionality to enable children to build their worlds. There are currently no proprietary tools available to do this. Several solutions to this problem are currently being investigated including using ActiveWorlds Bot technology to develop object construction assistants and using Macromedia Director 8.5's 3D authoring capabilities.

During the past two years the children working on VERTEX have achieved a remarkable amount, individually and collectively, with the software and hardware at hand. It is their work, very often beyond our expectations, which has allowed us to plan for the future developments and challenges of this project.

## References

- BERS, M. 1999. Zora: a Graphical Multi-user Environment to Share Stories about the Self. In *Proceedings of Computer Support for Collaborative Learning (CSCL) 1999* pp 33-40. Lawrence Erlbaum Associates.
- JOHNSON, A., MOHER, T., AND OHLSSON, S. 1999. The Round Earth Project - Collaborative VR for Elementary School Kids. In *Proceedings of ACM SIGGRAPH 1999 Conference Abstracts and Applications 1999*, pp. 90-93.
- PAPERT, S. 1980. *Mindstorms: Children Computers and Powerful Ideas*. Basic Books, New York.
- ROUSSOS, M., JOHNSON, A., LEIGH, J., VASILAKIS, C., BARNES, C., AND MOHER, T. 1997. NICE. Combining Constructionism, Narrative and Collaboration in a Virtual Learning Environment. In *Proceedings of ACM SIGGRAPH 1997 Computer Graphics Volume 31 No.3* pp. 62 - 63.
- VYGOTSKY, L. S. 1978. *Mind in Society: The development of higher psychological processes*. Cambridge, Mass: Harvard University Press.



# Workshop: Building an Affordable Projective, Immersive Display

Dave Pape  
Res Umbrae  
dave.pape@acm.org

Josephine Anstey  
University at Buffalo  
jranstey@buffalo.edu

## Abstract

Immersive, interactive virtual reality is a tool with hypothetically limitless uses. However, so far it has been put to serious use primarily in technical application areas such as computational science, automotive engineering, and chemical exploration. Groups working in these fields often have large budgets and can afford expensive, advanced displays. VR should also be of value to schools and museums, but most of them have much smaller budgets than major research labs, or are not able to support high-end graphics workstations. A simple, affordable, projection based display system can make VR far more accessible. In schools, displays could be put into individual classrooms and not just a central computer lab. In the museum world, small institutions would be capable of showing cutting edge digital work that previously has been restricted to a few large museums.

This workshop describes the construction of a single screen, passive stereo, VR display based on commodity, or otherwise low-cost, components. There are many options available for the major elements of such a system and the basic system can be modified or adapted to many different styles of use. Figure 1 shows a photo of such a system in use at the University at Buffalo.

## 1 Construction

The core elements of the system are three PCs, two projectors, tracking, and polarized stereo. The main computer, for the graphics, is a dual processor Linux PC with a two-channel 3D video card. The tracking system is a PC with an electromagnetic, 6 degree-of-freedom position tracker and a wand interactive device. The audio system is a PC with a generic sound card and speakers (and optionally a mixer and more expensive speakers). The stereo display uses circularly polarizing filters for the two projectors and inexpensive polarized glasses. The screen used is one specifically chosen to preserve polarization in rear-projection, minimizing the amount of stereo crosstalk (e.g., ghosting). Table 1 lists the components and estimated prices for a typical system.

A number of design decisions went into the choice of each of the system components. We favor projection-based display over head-mounted displays due to the intended use in museums and



Figure 1 The projection display at the University of Buffalo.

schools. Projection displays are easier to use for people inexperienced with VR and are able to support whole groups of viewers. We use polarized stereo, rather than active stereo, for two major reasons—the glasses are significantly cheaper and less fragile than LCD shutter glasses and low-cost LCD and DLP projectors are not capable of the scan rates necessary for active stereo. The rear-projection display requires more floor space than a front-projection display, but we accept that cost because we want to allow people to get up close to the screen to interact directly with the virtual environment. Finally, we chose the Linux operating system for greater compatibility with existing VR tools and applications that we use, most of which were developed for SGI workstations. In other environments, other operating systems would work just as well.

Table 1 Typical budget for a tracked, passive stereo system.

Dual-processor Linux PC	\$ 2,500
Two-channel 3D graphics card	\$ 300
2 DLP projectors	\$ 7,000
Circularly polarizing filters	\$ 150
Polarization-preserving rear projection screen	\$ 1,500
Polarized glasses	\$ 200
Tracker PC	\$ 500
Electromagnetic tracker	\$ 1,500
Wand	\$ 2,000
Audio PC	\$ 1,000
Speakers	\$ 1,000
<b>Total</b>	<b>\$ 17,650</b>

## 2 Conclusion

In basic performance tests, the low-cost system is comparable to one using an SGI Onyx. The quality of the graphics is not quite up to the level of high-end systems, though—aliasing can be particularly bad with some PC graphics cards and most cards only perform at their best when using limited color depths and texture resolutions. Affordable LCD and DLP projectors also have lower resolution than high cost CRT projectors, but are superior in terms of brightness and contrast.

Our system is being used in undergraduate and graduate media studies courses, allowing students to program their own interactive virtual environments. The students generally have backgrounds in the arts, rather than computer science, and their computer experience is primarily with packaged software on PCs and Macintoshes. Because it is based on common PC hardware, the students involved in the courses have also been able to help in constructing and maintaining the VR system. They can move between their own computers and the VR system more easily than if it were a high-end workstation. In addition to the VR classes, the display has been used as an Internet videoconferencing tool, for a study group distributed between Buffalo, Chicago, and Bloomington, Indiana.

# Workshop: Life Drawing and 3D Figure Modeling with MAYA

Gregory P. Garvey  
Department of Computer Science and Interactive Digital Design  
Quinnipiac University

## Abstract

This workshop, first given at SIGGRAPH 2001 [Garvey 2002] is now offered in two-parts. Part I introduces the process of transferring life drawings into 3D models using MAYA followed by a life drawing session. In Part II, workshop attendees working in the CAL import digitized life drawings into MAYA for setup as image planes to guide the modeling process. The goal of the workshop is to explore and develop skills of observational figure drawing and integrate them with the process of 3D modeling.

## 1 From Inspiration to Practice

This workshop was inspired in part by the drawing exercises found in the book *The Natural Way to Draw*, by Kimon Nicolaides. Through a series of exercises that focus on contour, gesture, weight, modeling, memory, analysis of contrasting curved and straight lines, Nicolaides strove for a deeper understanding and empathy with the subject in learning how to represent it: "You should draw, not what the thing looks like, not even what it is, but what it is doing." [Nicolaides 1941]. *The Natural Way to Draw* is really a guide to learning how to observe the world of existence with all of one's faculties, senses, intelligence and creativity. It is essential training for anyone involved in arts of visualization.

In Part II, attendees use the simple technique of lofting to model the human figure. With this approach an arm, leg or even torso can be easily modeled by creating a series of "ribs" made of geometric primitives such as circles (Figure 1). Attendees are provided with sample tutorial drawings representing front and side views of a leg and an arm. In MAYA a digitized figure

drawing is imported and setup as an image plane. The 2D image plane serves as a reference to adjust the size of each circle from both the front and side orthogonal views. A lofted surface will follow the form created by the frame of circles. This part of the workshop was inspired by a short exercise for modeling an arm by lofting a series of circles introduced in *Mastering MAYA 2* [Harovas et al. 2000].

## 2 Life Drawing Session

Part I of the workshop focuses on the preparation and execution of life drawings. The drawing of limbs is emphasized because of the complexity of drawing and modeling the geometries of the hands, feet, or head. Two orthogonal views, capturing the contours of a pose should be prepared in order to create a proportionally correct model of a subject. Graph paper can be used to facilitate alignment and size. Major topological features of the anatomy clearly visible in the front and side views can also aid alignment. Drawings can be realistic, stylized or even abstract using different media such as charcoal, pen and ink, conte crayon or pencil.

With certain poses some orthogonal views do not align (Figure 2). Multiple non-orthogonal views can also be prepared and suggest many possibilities for further experiments in purposeful misalignment, distortion or even abstraction. Combining these different perspectives in a cubist-like fashion creates surface geometries very difficult to anticipate or visualize

through any other means. Experimentation with multiple views lends itself to an iterative, trial-and-error modeling process similar to activity of drawing.

Combining these modeling experiments with new advances in non-photo realistic rendering techniques points to unexplored creative possibilities inherent in the medium of 3D computer graphics.

## References

- GARVEY,G.2002. Life Drawing and 3D Figure Modeling with MAYA. *Leonardo* 35, 3. Excerpts reprinted courtesy of the MIT Press.
- HAROVAS, P., KUNDERT-GIBBS, J. AND LEE, P. 2000. *Mastering MAYA Complete 2*. SYBEX, Alameda, CA.
- NICOLAIDES, K. 1941. *The Natural Way to Draw*. Houghton Mifflin Company, Boston.p.15.

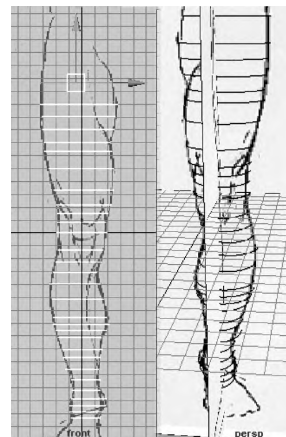


Figure 1 Circle are moved, scaled and aligned in the front and side views; completed circles are seen in the perspective view.

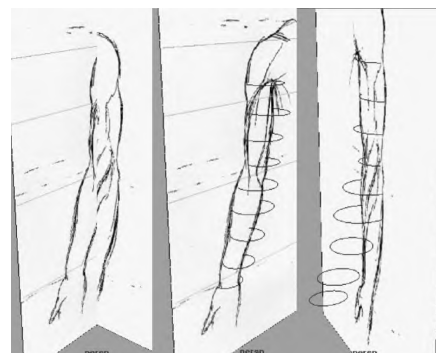


Figure 2 The front-and side-view orthogonal drawings of the arm in supination do not align.



# Workshop: Integrating Web 3D into 3D Animation Curricula

Mitch Williams  
3D-Online  
MWilliams@3D-Online.com

## Abstract

3D curriculum, once focused on modeling, rendering and animation, now includes web 3D and interactivity. But what options exist for new courses in web 3D? How in-depth are some web 3D tools and their learning curves? Should the curriculum expand to make 3D artists into programmers? What considerations should we give to the end-users' experience in viewing an online web 3D portfolio? This session examines how these issues are being addressed in teaching "Interactive Web 3D Media" at UC Berkeley, UC Irvine and UCLA Digital Arts and Entertainment Studies Extension. The session also provides hands-on experience with various tools and technologies.

## 1 Web 3D Technologies and Solutions Overview

There are some 40 web 3D technologies. Most import 3D models, add some interactivity via drag-and-drop interfaces or have full-fledged application programming interfaces (API's). But unlike Flash for 2D, there is no dominate web 3D tool. Each must be evaluated on its ease-of-use, adoption in the marketplace, costs, and end-user experience.

While teaching web 3D, the goals set for our courses were:

- Enable students to create and upload web 3D in their first class to help inspire and motivate them.
- Post web 3D content on a student's home page to demonstrate and showcase its use in web design.
- Integrate with existing 3D modeling programs such as 3D Studio Max, thus building on a student's skills or as an introduction to 3D.
- Consider end-user experiences in downloading plug-ins and performance in low-bandwidth.
- Minimize development tool and licensing costs.
- Teach programming as an advanced topic, preferring open, multi-purpose languages like JavaScript.

## 2 Solutions Fit for the Market Place

While also teaching "3D Graphics Programming" within Engineering Departments, we developed software tools for our art courses, allowing 3D artists to open and close doors; stop, start and reverse animations; and navigate freely or to specific camera positions. This tool, "Dynamic-3D", available for free, uses a Java-based, no plug-in, 3D engine. Students create their interactive web 3D within their 3D modeling program simply by appending keywords to the names of their 3D objects, exporting the 3D scene and coupling it with Dynamic-3D. Students can also rotate their 3D objects; turn lights on and off; and modify object colors. The applications for Web 3D include e-commerce, virtual worlds, web site design and product demonstrations.

This workshop will demonstrate how to develop interactive web 3D using Dynamic-3D without requiring students to learn any programming, nor download any plug-ins to create their portfolios. Attendees to this workshop will be shown how to go from exporting content from 3D Studio Max to creating various interactive web 3D applications.

## 3 Conclusions

Interactive Web 3D combines 3D modeling with software. Its creation relies on development tools of varying ease-of-use, costs and market acceptance. Yet Web 3D expands a student's skill set to future opportunities. Quickly creating and publishing an interactive 3D online portfolio motivates students to learn more and allows instructors to focus on scene optimization, design, performance, greater interactivity through programming and the end-user experience for web 3D.

# Workshop: Hi Tech – Lo Tech: K-12 Science Visualization

Susana Maria Halpine  
Candle Light Productions  
shalpine@earthlink.net

## Abstract

Integrating multimedia applications in the classroom can be overwhelming. Grants may address the cost of computer hardware, but where can instructors find the time to explore available software? Many visualization programs are free or low cost, but students will not grasp the importance of what they are viewing without proper conceptual introduction. Furthermore, many K-12 instructors are now expected to teach topics, including basic chemistry concepts, in which they may lack proper training.

The STArt! teaching Science Through Art program was developed to help teachers prepare for these educational challenges. Using an “Artist in Residence” format, workshops are developed in collaboration with participating teachers. Specifically, STArt! focuses on basic concepts addressed in the new California K-12 Science Content Standards. The program introduces molecular visualization software using narrative discussions, educational animation, and hands-on workshops using art materials and everyday objects. By exploring different learning modes, it makes basic science concepts more understandable to a broader audience. Furthermore, by collaborating with instructors within their classrooms, the program provides a creative resource for teachers in meeting the academic standards.

## 1 Technology Reuniting Science and Art

Despite its cultural importance, art is often considered expendable in grade school education. Art is among the first subjects to go with budget cuts, for example. Even more disheartening, in contrast to our increased reliance on technology, scientific literacy in America has been on a thirty year decline. Is there any wonder students have trouble “thinking visually” about science concepts?

An underlying principle of the STArt! program is that not only are art and science key elements for understanding the world around us, but the two are intimately related. Artists have intuitively used scientific methods, such as testing one variable at a time and noting the results, to develop artistic techniques with strong visual impact. Similarly, scientists use visual graphics and hands-on models to help conceptualize their research. The helical form of the DNA molecule, for example—whether represented through illustrations, sculptural models, or visualization software—is critical to understanding its genetic function.

Technology is reestablishing the relationship between the visual arts and science. In fact, some biochemistry professors are now requiring that their students be proficient in computer graphics.

## 2 How Many Atoms in the Entire World?

STArt! creates a model for integrating art and science education in a way that’s exciting, meaningful and accurate. Molecular models of familiar substances, such as colors and smells, are explored using the visualization freeware, WebLab Viewer Lite.

While presenting the STArt! program to a third grade class one student asked, “Wow, if there are 55 carbon atoms in one chlorophyll molecule, how many atoms are there in the entire

world?” After just one session viewing molecular models, the student was already applying the information to the world around him. Children’s curiosity and hunger for knowledge is boundless. With a creative approach to visualization software, students will begin to understand the scientific context of their everyday lives.



**Figure 1** The STArt! program integrates art and molecular visualization software in K-12 science education.

## References

- EGAN, K. 1992. *Imagination in Teaching and Learning*, The Middle School Years. Univ. of Chicago Press, 67–89.
- GLENN, J. AND COMMISSION. 2000. Before It’s Too Late: A Report to the Nation from the National Commission on Mathematics and Science Teaching for the 21st Century, The Middle School Years. U.S. Dept. of Education, 7.
- HALPINE, S. M. 2001. Molecular Visualization, a Microcosm of the E–Revolution. *IEEE MultiMedia* April–June, 4–7.
- HALPINE, S. M. 2001. Science Visualization and Educational Animation at SIGGRAPH 2001: The Next Big Deal. *Animation World Magazine*: [http://mag.awn.com/index.php3?ltype=pageone&article no=855](http://mag.awn.com/index.php3?ltype=pageone&article%no=855).
- PAPERT, S. 1980. *Mindstorms, Children, Computers and Powerful Ideas*. Basic Books, Inc., 38–54.
- RICH, M. Mathmol K-12 math and science visualization website. <http://www.nyu.edu/pages/mathmol>.
- SADKET, M. AND SADKER, D. 1994. *Failing at Fairness: How Our Schools Cheat Girls*. Touchstone/Simon & Schuster, 123.
- SILVER, H. F., STRONG, R. AND PERINI, M. J. 2000. *So Each May Learn, Integrating Learning Styles And Multiple Intelligences*. Association for Supervision and Curriculum Development.

# Workshop: Teaching Human Facial Modeling Through Plaster Face Casting

Adam Watkins  
University of the Incarnate Word  
watkinsw@universe.uiwtx.edu

## Abstract

As 3D student skills progress, sooner or later they begin work on accurately proportioned human models. While caricature design is often very forgiving in its realization, human designs need to be very close to “right on.” This presents a tremendous challenge for 3D animation students, many of whom have had limited anatomy training.

The traditional solution to source material has been to provide photographic references. The standard front and side shots provide a good starting point for students to work from but provide no information on appropriate polygon topology or details such as the curvature of the head between the eye and the temple. Yet, it is impractical to get a live model to provide 3D reference by sitting next to the student as he or she models. The difficulty of finding a good 3D reference for students to work from becomes the challenge.

At the University of the Incarnate Word, we have begun using a traditional method of face casting. Traditionally, this technique is used to create plaster masks or molds upon which prosthetics can be constructed. The mold is reusable and can be constructed from plaster or even lighter cements.

## 1 The Mold Creation Process

The student must first be prepared by covering her hair with either a skullcap or other thin plastic covering like a plastic wrap. A thin layer of petroleum jelly should be applied to eyebrows, eyelashes, and around the edge of face to ensure easy extraction from the alginate. To protect the clothes, cut a hole in a plastic bag and place it over the student’s shoulders and another across the student’s lap.

The technique utilizes an inexpensive and easy to use prosthetic alginate that is spread over the student’s face. For most individuals, about one-half pound is sufficient to completely cover the face, bottom of the chin and top of the neck. Be sure that in the application process you leave the nostrils uncovered.

The alginate sets quickly (within 5 minutes) and then is covered with plaster bandages to reinforce the form. Upon drying, the alginate and plaster bandages create a reinforced negative of the student’s face. It is important that the student does not move, talk, laugh, or smile during the process. Typically it takes two people to apply the alginate quickly enough to cover the face before it sets.

Once dry, have the student place her hands on her face and lean forward. Remove the alginate plaster form from the student’s face by placing the fingers inside the edge of the mask and working along the edge until the mask falls away from the student’s face.

Fill a box with loosely crumpled newspaper to create a soft place in which to place the alginate negative. Center the negative in the box in a way to allow the most plaster to be placed in the form. Fill the nostril holes with excess alginate or petroleum jelly to keep plaster from flowing through these holes. If there are any holes or bubbles that emerged during the process of applying the

alginate, take a moment to fill these imperfections with a new batch of alginate or petroleum jelly. Remember that holes in alginate mold will result in raised areas in the plaster form.

The best results are usually achieved by mixing the plaster to a fairly thick consistency. However, don’t use too much water. Filling this negative with Plaster of Paris or cement patching materials provides an accurate (to the pore) duplication of the student’s face. Typically, the drying process is under an hour.

When the plaster is dry simply remove the plaster and alginate negative from the plaster leaving the positive face form. Often, there are small imperfections in the positive form that can be easily worked out or sanded down with an Exacto knife, razor or sandpaper.

## 2 The Digital Realization

In class, we use this 3D plaster mask to discuss and draw appropriate polygon or NURBS topology before starting the digital modeling process. The students are assigned to begin plotting their polygon or curves placement by drawing directly on the face. Many digital modeling problems are avoided at this stage by making sure that the student plans out what parts of the face need the most digital information. Insightful discussion can take place as the student draws and erases planned topology on the face. This keeps the frustration level within the software low.

Upon completion of plotting the topology, the face mold can either be photographed from the side and front for use as the photo source or the student can work from the traditional photographs of their own face. However, once the topology is already worked out on the face, the student’s photographs of these molds often provide helpful information as the student begins to actually place curves or polygons. As they model using the photos as background reference, they can reference their own plaster face sitting on the desk in front of them.

For difficult-to-model areas of the face including the cheeks and the corners of the face (i.e., from the eye to the temple), the plaster mold provides invaluable reference unavailable from any of the front or side photographs. Other areas of specific benefit include the nostrils and bottom of the chin.

## 3 In Class Advantages

Typically, most 3D students are of the mindset that they should always be working on the computer. The technique described in this contribution is one way of assisting students to develop a broader sense of research and careful observation before, during, and after the digital creation process.

By creating a 3D replication of a student’s face, we have found an increase in productivity and a decrease in frustration. Theory of how to organize polygons for creases, folds, and bumps is understood before digital modeling takes place. Students have an opportunity to enjoy a tactile version of the product they are working on and generally enjoy a higher interest level.

# Emerging Technologies – Introduction

“...at its heart, the quest of SIGGRAPH has been to blaze the trail towards the seamless integration of man and machine”

e•merge (i-murj) v. (e•merged, e•merg•ing)

1. to come up or out into view.
2. (of fact or ideas) to be revealed by investigation, to become obvious.

Emerging Technologies is about extending the boundaries of how we consider the integration of man and machine. On one hand, the interface between man and machine has never been clean. All the research presented is about arcing the gap at that interface. On the other hand, our map of interactive techniques is constantly being refined but continues to show spans of the unknown. For 2002, we not only present the point on the map where we are today, but we have also focused on the paths forward that will fill in the spaces between the technologies.

Emerging Technologies 2002 brings you research from points around the world as well as points around the interactive techniques map. From robots in the physical world to humans in a virtual world and augmented worlds in between, from new display technologies to new input devices, including audio and haptics, the program gives you the opportunity to chart you own expedition.

I hope you will help us continue to define, estimate, and chart the edges of human/machine integration and that the reference point we provide will inspire your own vision and explorations.

Scott Senften, Chair

# Emerging Technologies – Committee

## Chair

Scott Senften  
SGI

## Committee

Preston Smith  
EDS

Ralph Loos  
University of Massachusetts - Amherst

David Newton  
University of Toronto

Kelissa Kesler  
Pixel Fuel Designs

Barb Helfer  
Helfer and Associates

John Fujii  
Hewlett-Packard Company

## Jury

Mk Haley  
Walt Disney Imagineering

Josh Strickon  
MIT Media Lab

Turner Whitted  
Microsoft Research

Chuck Hansen  
University of Utah

Joe Paradiso  
MIT Media Lab

Pete Tinker  
HRL Labs, LLC

Allen Bierbaum  
Iowa State University

## Advisory Board

Alan Commike  
QED Labs

Richard Marks  
Sony Computer Entertainment America

John Fujii  
Hewlett-Packard Company

Clark Dodsworth  
Osage Associates

Andrew Glassner  
Writer and Consultant

## Program Coordinator

Katie Rylander  
Capstone Solutions, Inc.



# Tomorrow's Yesterday: Mapping the E-Tech Continuum

John M. Fujii  
Hewlett-Packard Company \*  
fujii@siggraph.org

In computer graphics, history frames the possible, imagination paints the impossible, and passion fills in the rest.

Visionaries, explorers, pioneers, artists, and philosophers have shaped the Emerging Technologies programs since 1991. This history-mapping project recognizes their contributions and provides a starting point to envision the potentials of “what is next” and “who might take us there.”

The project idea grew from inspirations by ancient explorers' maps. Maps elegantly tell the story of exploration by revealing the fluid edges of “the known” and “unknown” and providing a visual context for “where to explore next.” They are a framework for viewing space and time.

The continuum of Emerging Technologies programs presents a similarly rich set of frontiers to explore. Where oceans and continents challenged the explorer on the move, the historian is faced with variable or incomplete records, uneven timelines, evolving technology and practices, and overwhelming volumes of information to study and digest.

Eleven years of history were scattered in a variety of formats and content. Gaps and inconsistencies that were encountered had to be supplemented by vast searches of paper reports, brochures, clippings, email archives, Internet searches, and dialogs with original contributors. Experience now suggests that we need to more actively prioritize preservation steps before the record becomes untraceable.

Challenges aside, the final results reveal a compelling look at evolution and revolution in action. The data follow journeys of now familiar themes such as hypertext, virtual reality, visualization, distributed interaction, simulation, avatars, spatial sound, autonomous agents, communities, telepresence, machine recognition, tangible interfaces, and immersive sensory feedback. From emergence to acceptance to integration into newer ideas, the creative continuum spreads like manifest destiny for those who perfect and improve the human / computer collaboration in the next century.

We believe that this project will be an ongoing community effort to appreciate and build upon SIGGRAPH's heritage. Like the maps of yesteryear, this work preserves the interesting artifacts of time (stale contact information, incomplete or missing descriptions, misinterpretations of the record, etc.) that will probably merge and yield to refinement in future “maps.” Like SIGGRAPH, it represents the state of what we know now, today, as it slips into the future to become Tomorrow's Yesterday.

Our understanding is just beginning...

To read further analysis and explore the research data visit:  
<http://www.siggraph.org/~fujii/etech/history.html>

\* Technical Computing Center, 3404 East Harmony Road, Fort Collins, CO 80528-9599



Detail of *Americae sive quartae orbis partis nova et exactissima descriptio*, Diego Gutiérrez [1562]

Library of Congress, Geography and Map Division  
Washington DC

<http://memory.loc.gov/ammem/gmdhtml/gutierrz.html>

# A New Step-in-Place Locomotion Interface for Virtual Environment With Large Display System

Laroussi Bouguila\*, Makoto Sato\*, Shoichi Hasegawa\*, Hashimoto Naoki\*,  
Naoki Matsumoto\*, Atsushi Toyama\*, Jelel Ezzine\*\*, Dalel Maghrebi\*\*  
\*Tokyo Institute of Technology - Precision and Intelligence Laboratory  
\*\*Ecole Nationale d'Ingenieurs de Tunis -  
laroussi@hi.pi.titech.ac.jp

## Abstract

The project presents a new locomotion interface for virtual environment with large display system. Users will be able to direct and control the traveling in the VE by in-place stepping and turning actions. Using a turntable technology, Visual feedback is continuously provided through the use of screen of limited size.

## 1. Introduction

Walking or traveling in general is an important life enhancing activity. It is initiated and sustained by the lower part of human body and considered as a necessary daily behavior that human performs to get from place to place. Keeping the same mean of active and dynamic ability to move through large-scale virtual environment (VE) will be of great interest for many applications demanding locomotion, such as building evaluation, urban planning, terrain exploration, and military and vocational training. In the present study we present a new virtual locomotion system that can provide users with the ability to walk through large-scale VE in an active "life-like" manner.

## 2. System Overview

The developed system as illustrated in figure 1, is composed of three main parts: a walking platform, sensing system, and large display system. The interface employs a turntable as walking platform on top of which users will stand and interact with the virtual environment that being projected on the large screen. Initially users will stand at the center of the turntable and face the center of the large screen. Users can engage into a virtual walking experience by stepping in place without propelling their body. Step-in-place movement, which is detected by a set of sensors embedded in the turntable, is treated as a gesture of moving forward in the VE. To change the moving direction or to explore the surrounding environment in general, users are required to turn their body about its vertical axis while remaining at the same position, the same natural turning action they perform in real life. The turn-in-place action is treated as a gesture changing the direction of the viewpoint. However, as large screen provides a limited projection

area, a large turning action will put the displayed image outside users' visual field of view. To overcome such limitation and keep users continuously oriented toward the screen and provided with enough visual display, the turntable platform will smoothly and passively rotate in the opposite direction of users turning. The passive compensation of users' active turning will continue until users regain their initial orientation.

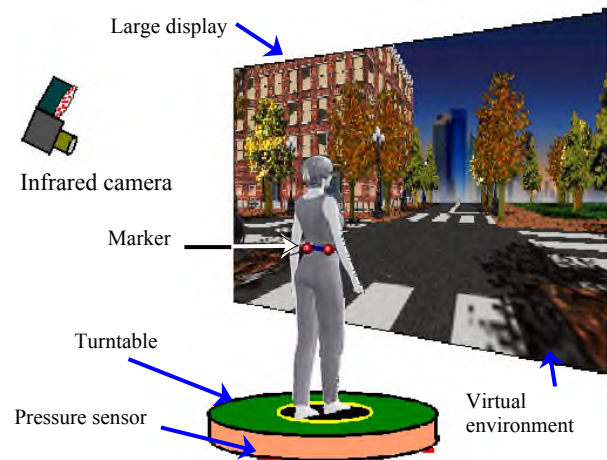


Figure 1: An overview of the proposed system.

## 3. Conclusion and Future Work

Locomotion is initiated and sustained by the lower part of the body as in real life. Virtual displacement is achieved by step-in-place and side-step actions, whereas a natural rotation about the vertical axis of the body controls virtual turning. Such physical based and life-like actions allow users to keep their natural reflexes and navigational skills to move around VE. The three distinguishing advantages of such technique are: 1) the use of turntable technique as walking platform to cancel turning rather than displacement 2) allow virtual surrounding projection with a single large screen. Users can make full turns in the VE without losing sight of the projected image. 3) Finally, The system demands no wired attachment to the user's body.

Farther study are necessary to evaluate the effect of turntable system on users' spatial perception and way finding capability.



# ARS BOX with Palmist -

## Advanced VR-System based on commodity hardware

Horst Hörtnner  
Ars Electronica  
Futurelab  
[horst.hoertner@aec.at](mailto:horst.hoertner@aec.at)

Christopher Lindinger  
Ars Electronica  
Futurelab  
[christopherli@aec.at](mailto:christopherli@aec.at)

Robert Praxmarer  
Ars Electronica  
Futurelab  
[robertpr@aec.at](mailto:robertpr@aec.at)

Andreas Riedler  
Ars Electronica  
Futurelab  
[andreas.riedler@aec.at](mailto:andreas.riedler@aec.at)

### Abstract

The ARS BOX is a projection-based (cave-like), PC-based VR system. It significantly reduces the time and money needed to develop and present Immersive Virtual Environment (IVE) applications while simultaneously expanding the options available compared to similar systems. A handheld PC serves as its interaction interface, making possible numerous innovative applications.

**CR Categories:** B.4.2 (Input/Output and Data Communication) – Input/Output Devices, J.m (Computer Applications)

**Keywords:** Virtual Environment, Device, Interface, commodity hardware

### 1 Introduction

For several years, one of the primary focuses of the Ars Electronica Futurelab's work has been the development of VR systems based on low-cost gaming technology both in terms of software (e.g. the use of game engines) and hardware (e.g. the use of interface devices, graphics adapters, etc). In this context, a huge number of tools and products were tested, and requirements were established in order to develop, use and improve applications in the field of art, research and technology. The chief motivation for taking up this field of research has been the steadily growing demand for affordable VR solutions for specific tasks.

### 2 The Palmist

The Palmist represents the effort to exploit the possibilities created by the use of a tracked handheld PC with wireless network access instead of conventional interfaces. It offers wireless mobility for the user that, in our view, increases the immersion of the user. Plus, the 2-D paradigm is much better for presentation of data like values and long texts and increases the usability of IVEs. This raises many new prospects for interaction with VR devices and VR applications. Last but not least, the Palmist is an 'intelligent' device, which means it's a computer whose capabilities can be expanded and achieve the full functionality attainable by the software. Our software framework (FATE) utilizes functionality for the handheld PC interface in the fields of rapid prototyping, design-on-the-fly, advanced navigation mechanisms for vast IVEs, control of the IVE, new interaction possibilities for simulation and CAD applications. We have developed several demo applications in order to demonstrate the functionality of FATE.

### 3. Conclusion

Current IVE applications, especially in the fields of industry, simulation and architectural visualization, suffer from the restricted interaction possibilities afforded by ordinary haptic interaction devices such as data gloves, wands, etc. To overcome these deficiencies of ordinary interaction devices, we launched the Palmist project to come up with an 'intelligent' interface for the ARS BOX. The idea is to have an easy-to-use interface that offers expandable capabilities and doesn't decrease user mobility to ensure unencumbered immersion.

The general aim of our work is to facilitate access to VR applications for designers and users, and at the same time to pave the way for new or improved applications in a wide variety of fields. The presentation within the Emerging Technologies Section is meant to signal that elaborate VR systems are no longer necessarily associated with extremely high costs. Through the use of the Palmist interface, we aim to demonstrate that a combination of pre-existing consumer hardware and innovative (software) approaches can yield results that are in some instances true breakthroughs.



Navigation Tool -Prototype

# Augmenting the Reality with 3D Sound Sources

Michael Haller  
Fachhochschule Hagenberg  
(MTD)  
[haller@fh-hagenberg.at](mailto:haller@fh-hagenberg.at)

Daniel Dobler  
Fachhochschule Hagenberg  
(MTD)

Philipp Stampfl  
Fachhochschule Hagenberg  
(MTD)

## The Augmented Sound Reality Application

Augmented reality (AR) is not only a new type of computer entertainment, but it can also be used for serious applications. Because the user is directly involved in the virtual world, enhanced reality can be more engaging than traditional computer work. Especially for a better sound impression, AR could be a solution for many problems. The sound component is still missing in current AR applications, which combine live video and computer graphics to produce real-time visual effects.



Figure 1: The user can place different 3D sound sources in the real world.

The ASR (Augmented Sound Reality) application allows the user to place and move different sound sources in the real world. Using the Augmented Reality technology the user is equipped with see through displays and a simple pen as input device places the 3D sound sources in the real world. The different sound sources are represented by 3D objects on the HMD which can be placed freely by the user. The advantage of the ASR installation is the direct integration and testing possibility of 3D sound in an Augmented Reality environment. The user can directly manipulate the 3D sound sources and observe the results immediately. Another advantage of ASR is the low cost setup (PC, webcam, i-glasses).

## Future Potential

Current AR applications do not allow a direct manipulation of sound sources. ASR is the first prototype which combines sound and graphics in a AR environment and offers a wide spectrum for further applications. Especially in the authoring process for Augmented Reality applications it becomes difficult to place the sound sources using traditional 2D/3D based authoring tools. Now, the authors can place 3D sound sources in the real 3D space and they have a more intuitive experience how it really sounds. In the second application, which we are working on, the user can place virtual furniture in the room. This application will be combined with ASR, so that the user doesn't see the newly established room, but he also can hear how the radio or DVD player sounds.



Figure 2: The ASR application allows a drag and drop of sound sources.

<http://webster.fhs-hagenberg.ac.at/staff/haller/projects/ws0102/asr/index.html>

<http://webster.fhs-hagenberg.ac.at/staff/haller/projects/index.html>

# Audio Haptics

Hiroo Iwata

Hiroaki Yano

Hiromi Igawa

University of Tsukuba  
iwata@kz.tsukuba.ac.jp

## Abstract

Audio Haptics is a rendering technique for auditory and haptic sensations. Sound and force are generated by using a physical model of virtual objects. We developed a software for real-time calculation of a physical model of virtual objects. A speaker is set at the grip of the haptic interface for spatial localization of the sound.

**Keywords:** sound rendering haptics, physical model

## 1 Purposes and goals of the project

Auditory information plays an important role in interaction with virtual objects. When we hit an object in the real world, we can hear sound emitted from the object. We also feel a reaction force simultaneously. This natural phenomenon isn't realized in virtual environments. The typical technique for auditory feedback in virtual environments is modulation of pre-recorded sounds. However, generation of sound depends on the shape, attributes, inner structure and locations of impact. The pre-recorded technique can't be applied to such conditions.

The major goal of this project is generation of sound and reaction force from a physical model of a virtual object. A physical model enables the user to hear sounds of new shapes and material of virtual objects. The user can also hear sounds from arbitrary impact points of the virtual object by physical model.

## 2 Technical Innovation of the Project

We developed a software for real-time calculation of physical models of virtual objects. Audio rendering is composed of three parts: (1) collision detection, (2) calculation of the vibration of virtual objects, and (3) sound emission. Finite Element Method (FEM) is used to obtain vibration of the object. Sound emission is calculated using sound pressure value from a partial-differential-equation. A speaker is set at the grip of the haptic interface for spatial localization of the sound.

The physical model is also used for the haptic rendering. Reaction force is calculated according to shape, elasticity and viscosity of the virtual object. Sound and force output are displayed to the user simultaneously using shared memory. The calculation can be done in real-time using a 2GHz Pentium 4 PC.

## 3 Larger implications of the project beyond this demonstration phase

Potential applications of the Audio Haptics are:

### (1) Virtual instrument

A new musical instrument can be created using the Audio Haptics technology. For example, the user can change the attributes of material of percussion instrument and can create new sound.

### (2) Inspection simulator

People often hit objects when they inspect them. They can obtain knowledge about the physical characteristics of an object, such as density distribution, by hitting it and hearing the sound. A simulator for such inspection can be built using the Audio Haptics technology.

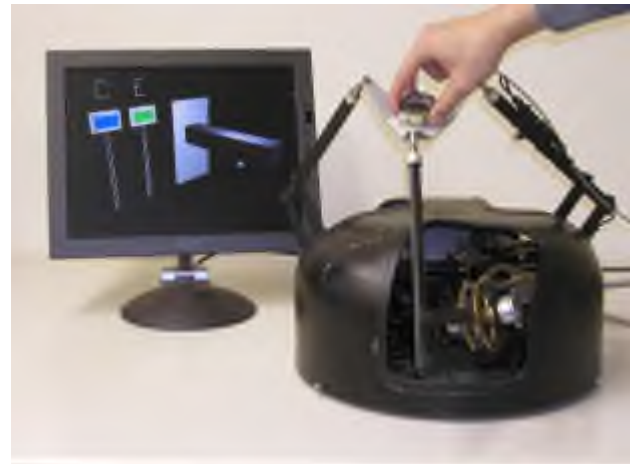


Figure 1. Overall view of the system

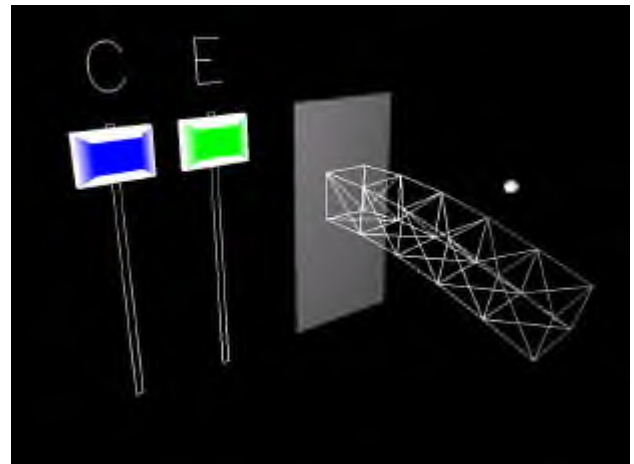


Figure 2. FEM model of virtual object

# Block Jam

Henry Newton-Dunn  
Sony CSL Interaction Lab  
henry@csl.sony.co.jp

Hiroaki Nakano  
Sony Design Center  
nakano@dc.sony.co.jp

James Gibson  
Sony Design Center Europe  
gibson@dc.sony.co.jp

## Abstract

Block Jam is a musical interface controlled by the arrangement of 24 tangible blocks. By positioning the blocks (Figure 1), musical phrases and sequences are created, allowing multiple users to play and collaborate.

**Keywords:** tangible interface, interactive music, collaborative systems



Figure 1: Photograph of the tangible interface (blocks)

## Overview

Block Jam takes advantage of both graphical and tangible user interfaces. Each block has a visual display and two modes of input, clickable and gestural. Each Block contains a sound group that can be chosen via the gestural input. The clickable input changes a block functionally (e.g. start or stop a sequence). Thus musically complex and engaging compositions can be rapidly assembled.

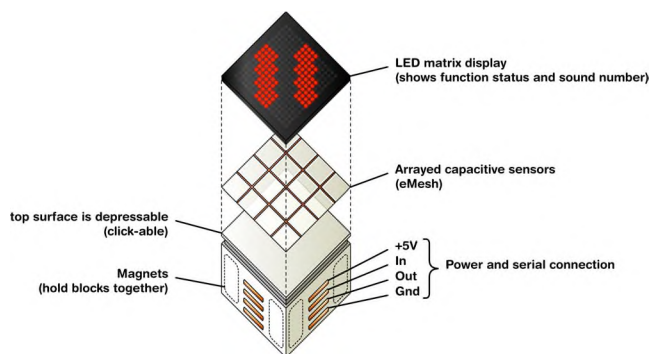


Figure 2, Anatomy of a block

The tangible nature of the blocks and the intuitive interface promotes face-to-face collaboration, and the presence of the GUI allows for remote collaboration across a network. By creating a combination of a tangible and a visual language, we are able to create endless meaningful musical structures in a novel and intuitive way.

## Goals

We believe in a future where music will no longer be considered a linear composition, but a dynamic structure, and musical composition will extend to interaction. We also believe that the divisions of composer, performer, and audience will be blurred, by the introduction of such media.

Our second aim is to put the group experience back into music. We understand that the musical experience changes with technology. Musical technology allows greater control and more possibilities than ever before and greater access to the beginner or novice. The way we receive and listen to music is also changing, not just in terms of Low-Fi to Hi-Fi or Phono to Tape to CD to MD to MP3, but also in terms of experience, music is moving from a social experience to a personal experience, from campfire to orchestra to living room to Walkman. Degrees of separation have occurred between the composer and the performer, the performer and the audience. This trend continues. Inversely, technology is moving towards community, towards the group, towards the network.

## Conclusion

Hence, we would suggest a need for a collaborative interactive music technology. We are not suggesting a musical toy or a novel musical instrument, but a discreet networked controller allowing multiple users the ability to control or modulate different elements of a single musical construct, to create a shared musical experience.

We envision a musical experience that can be shared equally by the novice or the musically adept, retaining a notion of author, of composer, yet allowing the user enough creative flexibility to add their own stamp. When a user's action is contextualized by the actions of other users, the experience will become shared, and perhaps be not so far from the campfire of before.

## References

- POUPYREV, I. "Augmented Groove: Collaborative Jamming in Augmented Reality." in SIGGRAPH 2000 Conference Abstracts and Applications, ACM Press, NY, p.77
- GORBET, M., ORTH, M., AND ISHII, H. "Triangles: Tangible Interface for Manipulation and Exploration of Digital Information Topography." Proceedings of SIGCHI1998: ACM Press
- JUN REKIMOTO. "SmartSkin: An Infrastructure for Freehand Manipulations on Interactive Surfaces." to be published in ACM SIGCHI2002

# Cyberarium Knowledge Fountain

Center for Really Neat Research

Cyberarium knowledge fountain is a opportunity:

- To demonstrate socially responsible applications of communication technology
- To provide a social exchange at which multiple communities can explore unconventional applications of advanced technological concepts
- To identify key areas where information technologies can be effectively applied to improve quality of life
- To encourage young minds to explore science and think about how to make the world a better place

The Cyberarium Knowledge Fountain is an innovative, integrative and interactive platform which will stimulate discovery and exploration of creative and socially enriching interactive techniques through intelligence engaging, imagination inducing hands-on experiences with novel human-computer interaction devices.

The knowledge fountain is a portable, self-contained, easily deployed workstation configurable for use as a rapid prototyping system for new human-computer interfaces. Using various sensors, interface boxes, and data-flow networking software, the knowledge fountain allows for rapid development of new ways to get information from the physical environment to the communication system. NeatViz software enhances the system by providing new ways to perceptualize complex data streams and information systems. Knowledge fountain is being used in schools and cyber camps as a tool for teaching kids about information systems.

NeatTools software is freely downloadable from our website [www.pulsar.org](http://www.pulsar.org), and the NeatViz software is freely downloadable from <http://ackbar.mindtel.com>.

Please visit our other websites for more information:

[www.cyberarium.com](http://www.cyberarium.com)

[www.mindtel.com](http://www.mindtel.com)

[www.medibolt.com](http://www.medibolt.com)

It is not enough to just refine current methods of education; instead, we must reach out with all our insight and an open mind and begin to develop new methods of thinking, learning, exploring and growing which will serve as the basis of intelligent migration into this new and exciting future. We must be proactive and make a conscious effort to act in ways which will allow us to reach the most potential offered by this inevitable progression into a new century and a new world. By developing/orchestrating a compassionate and pro-social basis for exploring and sharing knowledge about the impact of technology we can create an environment for intelligently cultivating an awareness of the processes which will shape the nature of the future. It becomes our obligation as leaders in the revolution of thought and process that we take steps to ensure that we have given our future generations the best possible opportunity to reach the highest levels of their true potential. We must overcome the natural fear of the unknown and move forward with determination and compassion as the future generations will indeed be strongly affected by what we do in these early phases of the information culture.

As we live in an information age, our lives are continually influenced by new applications of communication technology. Informatic systems impact society at all levels, from personal and interpersonal dynamics, to global systems of immense complexity. The future needs of the info-culture can be anticipated, thus environments must be created and processes need be initiated which will aid in a smooth and intelligent transition into the future. We must reach out with all our insight and an open mind and begin to develop new methods of thinking, learning, exploring and growing which will serve as the basis of intelligent migration into this new and exciting future.

Contact Information:

Dave Warner, MD, PhD  
[davew@mindtel.com](mailto:davew@mindtel.com)

Matt Carbone  
[matt@mindtel.com](mailto:matt@mindtel.com)

office: 252-752-5397  
fax: 252-752-7356



# Distributed Systems of Self-reconfiguring Robots

Zack Butler, Robert Fitch, Keith Kotay and Daniela Rus\*  
Department of Computer Science  
Dartmouth College

## 1 Introduction

A robot designed for a single purpose can perform a specific task very well, but it may perform poorly on a different task, or in a different environment. This is acceptable if the environment is structured; however, if the task is in an unknown environment, then a robot with the ability to change shape to suit the environment and the required functionality will be more likely to succeed. We wish to create more versatile robots by using self-reconfiguration: hundreds of small modules will autonomously organize and reorganize as geometric structures to best fit the terrain on which the robot has to move, the shape of the object the robot has to manipulate, or the sensing needs for the given task. For example, a robot could synthesize a snake shape to travel through a narrow tunnel, and then morph into a six-legged insect to navigate on rough terrain upon exit.

Self-reconfiguring robots are well-suited for tasks in hazardous and remote environments, especially when the environmental model and the task specifications are uncertain. A collection of simple, modular robots endowed with self-reconfiguration capabilities can conform to the shape of the terrain for locomotion by implementing “water-flow” like locomotion gaits which allow the robots to move by conforming to the shape of the terrain.

To create autonomous robot systems capable of such applications, our research agenda is focused on two directions of work: (1) new designs for modular robots that can support self-assembly and self-reconfiguration and (2) new distributed planners that support parallelism, are efficient, and correctly direct units to change shape.

## 2 Robots and Capabilities

Several groups have examined how to design robot modules capable of self-reconfiguration [Rus and Chirikjian 2001]. In our own work [Rus et al. 2002] we have developed two unit-modular systems. The main goal for the Molecule Robot has been to achieve self-reconfiguration in three dimensions. The Crystal Robot uses a novel actuation mechanism, scaling, which gives more robust motion than the previous rotation-based actuation systems.

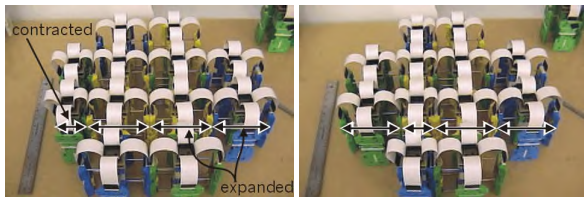


Figure 1: A 12-module Crystal robot performing locomotion.

The Crystal Robot (Figure 1) has some of the motive properties of muscles and can be closely packed in 3D space by attaching

\*e-mail: rus@cs.dartmouth.edu. Support for this work was provided through the NSF CAREER award IRI-9624286 and NSF awards IRI-9714332, EIA-9901589, IIS-9818299, and IIS-9912193, the Sloan Foundation, and a NASA Spacegrant award.

units to each other. Each module actuates by expansion and contraction. By expanding and contracting neighbors in a connected structure, an individual module can be moved in general ways relative to the entire structure. Crystal modules never rotate relative to each other; their movement is actuated by sliding via expansion/contraction. This basic operation leads to new algorithms for global self-reconfiguration planning. Each module has on-board processing, sensing, communication and power.

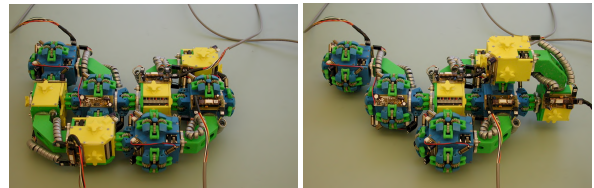


Figure 2: A Molecule Robot consisting of four modules can locomote using a “tumbling” gait.

A Molecule Robot [Kotay and Rus 1999] consists of multiple units called Molecules; each Molecule consists of two atoms linked by a rigid connection called a bond (see Figure 2). Each atom has five inter-Molecule connection points and two degrees of freedom. One degree of freedom allows the atom to rotate  $180^\circ$  relative to its bond connection, and the other degree of freedom allows the atom (thus the entire Molecule) to rotate  $180^\circ$  relative to one of the inter-Molecule connectors at a right angle to the bond connection. A feature of our prototype is the use of a robust gripper-type connection mechanism.

Crystal and Molecule robots are dynamic structures. They can move using sequences of reconfigurations to implement locomotion gaits, and they can undergo shape metamorphosis. We have developed several planning algorithms (both centralized and distributed) to support these capabilities [Rus et al. 2002; Kotay and Rus 1999].

## 3 Conclusion

Creating self-reconfiguring robots is a considerable challenge which can be met through new designs for reconfigurable systems and new ideas on algorithmic planning and control. Our robots are very encouraging first steps toward creating new robotics applications.

## References

- KOTAY, K., AND RUS, D. 1999. Locomotion versatility through self-reconfiguration. *Robotics and Autonomous Systems* 26, 217–232.
- RUS, D., AND CHIRIKJIAN, G. 2001. Special issue on self-reconfigurable robots. *Autonomous Robots* 10, 1, 1–124.
- RUS, D., BUTLER, Z., KOTAY, K., AND VONA, M. 2002. Self-reconfiguring robots. *Communications of the ACM* 45, 3, 39–45.

# Focus Plus Context Screens: Visual Context and Immersion on the Desktop

Patrick Baudisch  
Xerox PARC  
patrick.baudisch@acm.org

Nathan Good  
Xerox PARC  
ngood@acm.org

## Abstract

Focus plus context screens are wall-size, low-resolution displays with an embedded, high-resolution display region. Focus plus context screens allow users to view details of a document up close, while simultaneously seeing peripheral parts of the document in lower resolution. Application areas range from geographic information systems to interactive simulations and games.

## 1. Focus plus context screens

Many professional computer users, such as chip designers or geographic information specialists, work with visual documents that are too large and detailed to fit on their screens. Consequently, users have to access information sequentially using zooming and panning. *Focus plus context screens* are a new type of computer screen that addresses this issue. Focus plus context screens [Baudisch et al. 2001] are wall-size, low-res displays with an embedded hi-res display. Image content is displayed across the high-res and the low-res display regions, such that the scaling of the display content is preserved, while its resolution varies according to display region. Focus plus context screens allow users to view details of a document up close, while simultaneously seeing peripheral parts of the document in lower resolution.

Figure 1a shows the display prototype that we will demonstrate at SIGGRAPH. The frame holds a piece of foam core that serves as a projection surface (4x3 feet). A projector that is located behind the user projects onto it, turning the foam core into a large low-resolution display, the *context display*. The foam core has a hole in the center directly into which an LCD monitor is embedded, the hi-res *focus display*. Customized software is used to display image content across both display regions, such that the scaling of the image is preserved, while its resolution varies across the two display regions. Content panned into the focus region is viewed in higher detail. In Figure 1a, the focus region provides enough resolution to allow users to see individual cars on a satellite image of San Francisco. At the same time, the context region allows the user to stay oriented by providing a visual context containing important landmarks.

We have applied focus plus context screens to several application areas. On a focus plus context screen, editors of print products can work on details while constantly being aware of the detail's context, making it easier to judge how a detail modification will affect the overall impression (Figure 1b). Our adaptation of the Linux operating system allows running visual applications requiring many windows or large windows (Figure 1c). In a videoconference scenario, focus plus context screens allow simultaneous display of the remote partner *and* the detail information that this person presents, e.g. readable text (Figure 1d). Focus plus context screens increase the level of immersion of virtual reality applications, such as first person shooter games (Figure 1e). These games require players to be constantly focused on a cross-hair at the screen center and since human vision in the peripheral regions is limited to low-res, a focus plus context screen produces the same user experience as if the entire display was hi-res.

## 2. Conclusions

Focus plus context screens provide users with additional visual context that can be perceived using peripheral vision. This allows users to work faster and with fewer errors [Baudisch et al. 2002]. While focus plus context screens allow users to work effectively with large visual documents, they also increase the level of immersion in virtual worlds. Since only a portion of focus plus context screens is actually hi-res (similar to [Yoshida et al. 1995] with head-mounted displays), focus plus context screens are more than an order of magnitude less expensive than currently available projection array-based systems and also require less space [Hereld et al. 2000].



Figure 1: (a) Focus plus context screen prototype. The iconic illustration (bottom right) shows the location of the focus screen. The callout shows the different resolutions of focus and context area. Other application examples (b) Adobe Illustrator (c) Linux desktop (d) videoconferencing and (e) first person shooter game.

## References

- BAUDISCH, P., GOOD, N., BELLOTTI, V., SCHRAEDLEY, P.; Keeping Things in Context: A Comparative Evaluation of Focus Plus Context Screens, Overviews, and Zooming, to appear in *Proc. of CHI'02*.
- BAUDISCH, P., GOOD, N., STEWART, P. Focus plus context screens: Combining display technology with visualization techniques, *Proc. of UIST '01*, p. 31-40.
- HERELD, M., JUDSON, I., STEVENS, R., Introduction to Building Projection-based Tiled Display Systems *IEEE Computer Graphics & Applications* 20(4), 2000.
- YOSHIDA, A., ROLLAND, J. & REIF, J. (1995). Design and applications of a high resolution insert head-mounted-display, *Virtual Reality Annual International Symposium '95*, 84-93.



# Immersive and Interactive Rear-Projected Stereo DLP Reality Center

The Immersive and Interactive Rear-Projected Stereo DLP Reality Center aims to improve the quality of immersive visualization and increase the efficacy of seamless real-time interaction with complex stereoscopic data.

Front-projected displays can provide an effective immersive experience, but the ultimate value of such a display is constrained by the shadow of the user projected onto the screen, which occurs when the user is positioned close enough to the screen to interact with the data. This occlusion provides a constant and pervasive reminder to the user and audience that the data is indeed only a virtual representation of the real thing, thus lessening the ultimate experience of being immersed in the data, as well as diminishing the level of interactivity between the user and the data.

Conversely, the rear-projection of stereoscopic data onto a multi-channel curved-screen display can provide a profound immersive experience for both the user and the audience, especially when the user can seamlessly interact in real-time with the data.

Research and development activities in the stereoscopic projection industry in the past few years have focused largely on the Stereo DLP technology developed by Texas Instruments. Stereo DLP technology is becoming increasingly popular due to its providing an extremely bright high-contrast image, with minimal maintenance and re-alignment required.

However, the utilization of Stereo DLP technology for high-end multi-channel curved-screen displays, and particularly for those employing rear-projection, has not yet been fully developed commercially. This delay is due primarily to the technological barriers inherent in attempting to accurately geometrically warp and blend stereo imagery generated by the fixed-matrix Stereo DLP panels, which are at the core of the Stereo DLP projection technology, especially for rear-projection onto curved screens.

BARCO Simulation Products' extensive R&D in the field of Stereo DLP projection technology has resulted in the development of Warp 6, which uses sophisticated geometric image warping algorithms executed by processors internal to each projector. Technological advantages offered by Warp 6 technology derive primarily from the utilization of proprietary bicubic, rather than bilinear, interpolation algorithms.

BARCO has applied significant R&D resources towards solving the Stereo DLP image-warping challenges in its belief that multi-channel rear-projected curved-screen Stereo DLP display technology will ultimately solve problems that are currently associated with using front-projected curved-screen Stereo DLP displays, and also that it will provide a significant benefit over rear-projected flat-screen Stereo DLP displays.

The principle advantages of rear-projection curved-screen Stereo DLP display technology lie in its optimizing the potential immersive qualities of the display, and in providing better real-time interaction with the stereoscopic data being manipulated and visualized.

It is anticipated that this new projection technology will be relied upon heavily in the future. Further, its commercial use should actually foster the development of new software applications in many different fields, from scientific visualization to fine arts, which will take specific advantage of the improved immersive and interactive qualities of this type of display.

## Contact

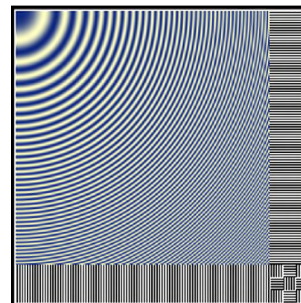
Andrew Joel  
BARCO Simulation Products  
600 Bellbrook Avenue  
Xenia, OH 45385 USA  
+1.770.218.3278  
+1.937.372.8645 fax  
Email: andrew.joel@barco.com

## Partners

SGI  
Digital ArtForms, Inc.

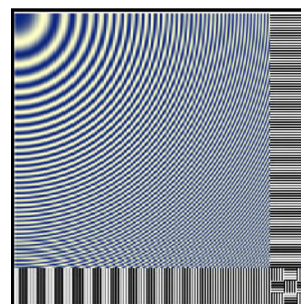
## Collaborators

Michael Brown, SGI  
Almos Elekes, SGI  
Paul Mlyniec, Digital ArtForms, Inc.  
Keith Tushingam, Schlumberger  
Malte Zoeckler, Indeed 3D



### Warp 6 Bicubic Interpolation

- 4 x 4 point
- Gamma correction preceding scaling
- Floating point accuracy



### Bilinear Interpolation

- 2 x 2 point
- Gamma correction after scaling
- 8 bit linear accuracy

# Lewis the Robotic Photographer

Cindy M. Grimm\*      William D. Smart†  
Department of Computer Science  
Washington University in St. Louis  
St. Louis, MO 63105  
United States

## 1 Introduction

Lewis is a (short) human-sized mobile robot that wanders through crowded rooms, taking pictures of people, much like a photographer at a wedding reception does. The goal is take high-quality, well-composed photographs of people non-intrusively, and to offer these pictures as keep-sakes of the conference.

## 2 The Exhibit

The exhibit space consists of an area for posed portrait shots and an open area where Lewis interacts with the public as they pass by the installation. Lewis will wander about this open area, looking for good pictures to take. Once a good picture opportunity is detected, Lewis will maneuver into place, take the picture and transmit it to a display workstation at the side of the exhibit. Subjects can then use this workstation to browse the pictures that Lewis has taken, choosing their favorites to be printed out as a keep-sake of their visit. In addition to simply printing out the pictures that were taken, users have the option of applying simple filters to the photographs to, for instance, have them rendered in the style of their favorite impressionist painter.

In addition to being a wandering photographer, Lewis can also take more formal portrait shots of individuals or groups. Subjects can sign up in advance for a formal sitting, where Lewis will take their portraits against a variety of scenic backdrops.

## 3 Robot-Human Interaction

One of the main goals of this project is to explore how humans interact with computers outside of the usual desktop setting. One of the first successful mobile robots to be deployed in a public setting was Minerva [Thrun et al. 2000], which gave pre-set tours in the Smithsonian Museum of American History in Washington, D.C. In addition to the traditional problem of navigation through a complex environment, Minerva also had to deal effectively with the general public. This general public, it turned out, was not particularly robot-friendly. For example, children tend to try climbing on the moving robot, and adults purposively block sensors or stand in the robot's path to see what it will do. Thus, dealing the a dynamic human-populated environment proved to be much more difficult than dealing with a static, mostly unpopulated one.

Lewis must also deal with the problems of interacting with humans who may not be accustomed to robots. However, the quality of this interaction is different from Minerva's, mostly because of their different purposes. As a photographer, Lewis must try to navigate through the environment and take pictures without getting in the way, or ideally, without being noticed. This is different from Minerva, whose purpose was to be a tour-guide, and thus to be the center of attention.

Lewis relies mostly on vocal requests to get a person's attention. It has been shown that using specific tones of voice (polite, mildly annoyed, *etc.*) increases the chances that humans will interact with the robot [Schulte et al. 1999]. It has also been our experience that directing the camera towards the subject as the robot utters a phrase greatly increases the quality of the interaction. We suspect that using emotive voice tones and moving the camera make the robot appear more like a living thing, allowing a human subject to more easily ascribe real intention to it. This, in turn, makes it easier for humans to interact with the robot, since they can use a human-to-human mode of interaction (which they have), rather than a human-to-intelligent-machine mode (which they likely do not have).

These techniques are used to move people if they are blocking the robot's path ("Please may I get through?"), convince people to have their picture taken ("May I take your photograph?") and to time shots ("Say cheese!"). During the installation, we plan to experiment with different specific methods of interaction, in order to determine which ones work best in which situations.

## 4 The Technology

Lewis uses a set of basic navigation routines to localize itself within the installation, and to safely navigate through crowds of people. Using a small number of special landmarks, the robot keeps track of its location, allowing it to go to a specific point in the installation, such as the portrait shot area, easily.

As it travels about the installation, Lewis continually takes pictures with its video camera and searches them for potential faces. Once a suitable candidate face (or faces) has been found, Lewis calculates how it has to move in order to take a well-framed shot of that person. This interfaces with the low-level navigation routines to safely move Lewis to the appropriate position.

Potential faces are chosen as candidates for a photograph based on a very simple set of composition rules that define what constitutes a "good" picture. Once we have identified such a good picture, moved into place and taken it, the image is sent to a workstation for further processing and display.

## References

- SCHULTE, J., ROSENBERG, C., AND THRUN, S. 1999. Spontaneous short-term interaction with mobile robots in public places. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*.
- THRUN, S., BEETZ, M., BENNEWITZ, M., BURGARD, W., CREMERS, A. B., DELLAERT, F., FOX, D., HÄHNEL, D., ROSENBERG, C., ROY, N., SCHULTE, J., AND SCHULZ, D. 2000. Probabilistic algorithms and the interactive museum tour-guide robot Minerva. *International Journal of Robotics Research* 19, 11, 972–999.

\*Email: cmg@cs.wustl.edu

†Email: wds@cs.wustl.edu

# NONA-Vision

Hiroo Iwata Hiroaki Yano Motohiro Tsuzuki Fumitaka Nakaizumi Takayuki Yoshioka Yutaka Miyakita  
University of Tsukuba  
iwata@kz.tsukuba.ac.jp

## Abstract

The NONA-Vision is a high-resolution, wide-angle video capture and projection system. The display is composed of nine rear-projection screens. Images for the nine screens are captured by a specialized camera-head, in which the optical centers of nine video cameras are located at the identical position.

**Keywords:** immersive projection display, telepresence, video image

## 1 Purposes and goals of the project

Visual immersion plays an important role in virtual environment or telepresence. A large screen can cover a wide-angle field of view. Such a screen needs a large number of pixels. Current projectors have good resolution, and graphics workstations have the ability to display real-time computer graphics. However, high-resolution images of the real world are difficult to display. An HDTV camera has the highest resolution for video image, although its pixels aren't sufficient to cover the full field of view of human vision. Moreover, wireless transmission of HDTV images isn't available in general. The NONA-Vision aims to achieve high-resolution video images of remote environments by using specialized multiple cameras.

## 2 Technical Innovation of the Project

The NONA-Vision is a high-resolution, wide-angle video capture and projection system. The display is composed of nine rear-projection screens. Images for the nine screens are captured by nine video cameras. Mirrors are used to place optical centers of the cameras at the same position. Optical centers of the nine cameras are located at a single point by using eight mirrors. This configuration provides the images on the screen without distortion. Video images of each camera are transmitted using a UHF transmitter.

The display is mounted on a 360 degree turn table. Rotational motion of the viewer is measured by a position sensor. The display rotates around the viewer according to his/her motion, so that it displays a full-surround image.

## 3 Larger implications of the project beyond this demonstration phase

Potential applications of the NONA-Vision are teleoperation and virtual travel. A wide-angle image is useful for remote robots. Scenes of an un-known environment are difficult to recognize by using the narrow field of view of a conventional camera-head. The NONA-Vision provides a good sense of presence of remote sites.



Figure 1. Nine rear-projection screens



Figure 2. Camera-head with nine video cameras

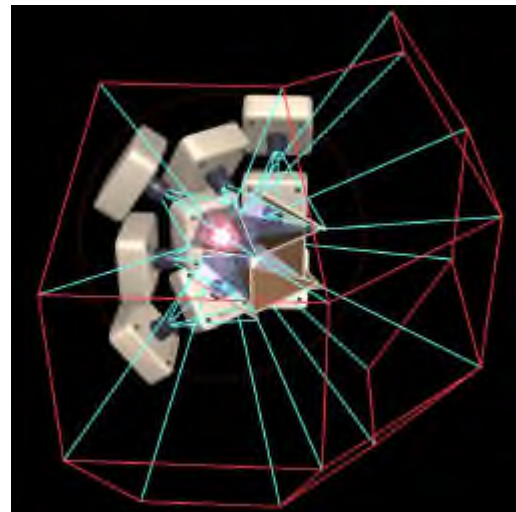


Figure 3. Optical center of the nine cameras

# Occlusive Optical See-through Displays in a Collaborative Setup

Kiyoshi Kiyokawa  
Communications Research Lab.  
kiyo@crl.go.jp

Hiroyuki Ohno  
Communications Research Lab.  
hohno@ohno-lab.org

Yoshinori Kurata  
Topcon Co.  
y.kurata@topcon.co.jp

## Introduction

There are several approaches to realize a 3D display that can be viewed by multiple co-located users, and each has its own pros and cons. A volumetric display utilizes afterimage effects, and it usually can only show transparent images within its volume. Some projection-based systems support independent viewpoints for more than two users [1], however, virtual objects can be shown only within a viewing frustum in front of the screen. On the other hand, head mount display (HMD) based Augmented Reality (AR) can show virtual objects at arbitrary locations [2]. Besides, an HMD can potentially show correct occlusion phenomena between virtual and real scenes. While a video see-through HMD severely degrades the quality of the real scene and adds a certain system delay, an optical see-through HMD keeps the intrinsic quality of the real scene. However, a virtual scene had to be a semi-transparent ghost due to a half-silvered optical combiner so far.

## ELMO: Our occlusive optical see-through HMD

Our optical see-through HMD, ELMO, is capable of showing opaque virtual objects, by using a secondary LCD to properly block incoming lights on a pixel basis [3]. Fig 1 shows a typical ghost image (left) and opaque virtual objects with correct occlusion attributes (right). ELMO realizes a 3D collaborative workspace that has the following three advantages.

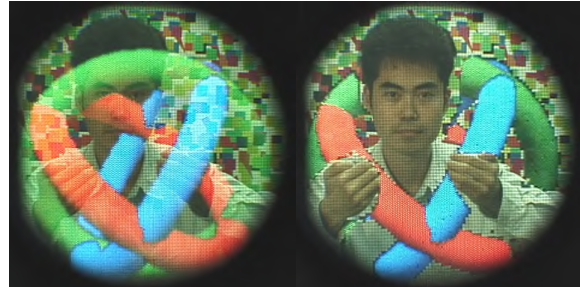
- 1) Unlike with a video see-through HMD, a user can see non-digitized real collaborators to have superior awareness and to receive natural non-verbal communication cues.
- 2) Unlike with a fixed display, virtual objects can be shown anywhere in the workspace, e.g., between two users' faces.
- 3) Unlike with a conventional optical see-through HMD, virtual objects can be transparent or opaque, as demanded by a certain task.

We have newly designed ELMO-4 that has lighter (1.5kg) and smaller optics, enough to wear, with absolutely no viewing offset (See Fig 2).

## Exhibition

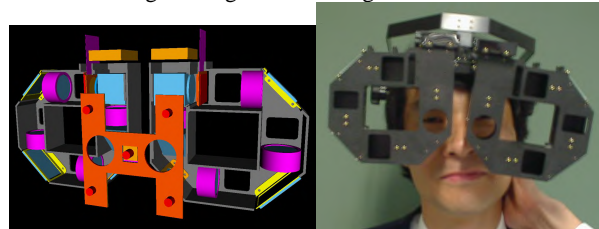
Fig 3 shows a layout of the exhibition. Users of this exhibition will enjoy a collaborative AR environment in which they can see realistic opaque virtual models in a non-digitized real world. In the environment, users will see virtual objects flying around them, which cannot be realized by a volumetric display or a projection-based display. They will also see a virtual object covering further partners' face and being covered by closer hands properly, which is impossible with a conventional optical see-through display. Users will notice the real world is naturally and clearly visible, as it is supposed to be, which will help them feel they co-exist in the real world, not in a video world.

Even though users still have to wear cumbersome devices, our collaborative environment has many important advantages for



(Left) Normal ghost image. (Right) Mutually occluded image.

Fig 1. Images seen through ELMO-3.



(Left) Optics design. (Right) Actual system (without cameras).

Fig 2. Appearances of ELMO-4.



Fig 3. Exhibition layout.

natural and flexible face-to-face 3D collaborative activities, which have never been combined all together before.

## References

1. Bimber, O., Frohlich, B., Schmalstieg, D., and Encarnacao, L.M., "The Virtual Showcase," *IEEE Computer Graphics & Applications*, vol. 21, no.6, pp. 48-55, 2001.
2. Schmalstieg, D., Fuhrmann, A., Szalavari, Z., Gervautz, M., "Studierstube - An Environment for Collaboration in Augmented Reality," *CVE '96 Workshop Proceedings*, 1996.
3. Kiyokawa, K., Kurata, Y., and Ohno, H., "An Optical See-through Display for Mutual Occlusion with a Real-time Stereo Vision System," *Elsevier Computer & Graphics, Special Issue on "Mixed Realities - Beyond Conventions"*, Vol.25, No.5, pp.2765-779, 2001.



# Physiological Reaction and Presence in Stressful Virtual Environments

Michael Meehan<sup>1,2</sup>, Mary Whitton<sup>1</sup>, Sharif Razzaque<sup>1</sup>, Paul Zimmons<sup>1</sup>, Brent Insko<sup>1</sup>, Greg Combe<sup>1</sup>, Ben Lok<sup>1</sup>, Thorsten Scheuermann<sup>1</sup>, Samir Naik<sup>1</sup>, Jason Jerald<sup>1</sup>, Mark Harris<sup>1</sup>, Angus Antley<sup>1</sup>, Frederick P. Brooks, Jr.<sup>1</sup>

<sup>1</sup>Department of Computer Science, University of North Carolina, Chapel Hill

<sup>2</sup>Department of Surgery, Stanford University

[meehan, whitton, sharif, zimmons, insko, coombe, lok, scheuerm, naiks, jjerard, harrism, antley, brooks]@cs.unc.edu

A common metric of VE quality is presence – the degree to which the user feels like they are *in* the virtual scene as opposed to the real world. Presence is important for many VE applications [Hodges *et al.* 1994]. Since presence is a subjective condition, it is most commonly measured by self-reporting, either during the VE experience or immediately afterwards by questionnaires. There is vigorous debate in the literature as to how to best measure presence [Meehan 2001].

In order to study a VE's effectiveness in evoking presence, researchers need a well-designed and verified measure of the phenomena. This exhibit demonstrates the compelling VE and physiological monitoring used to evaluate three physiological measures – heart rate, skin conductance, and skin temperature – as alternate measures of presence in stressful VEs. We found that among these physiological reactions, heart rate is the best measure of presence [Meehan *et al.* 2002].

Figure 1 shows the environment: a Training Room, quite ordinary, and an adjacent Pit Room, with an unguarded hole in the floor leading to a room 20 ft. below. On the upper level the Pit Room is bordered with a 2-foot wide walkway. The 18x32 foot, 2-room virtual space fits on a one-to-one scale within the SIGGRAPH Emerging Technologies exhibit and is underneath the 3<sup>rd</sup> Tech wide-area ceiling tracker. Users, wearing a head-tracked stereoscopic head-mounted display and physiological monitoring equipment (Figure 2), practice walking about and picking up and placing objects in the Training Room. Once the door between the two rooms opens, they walk to an unexpected hazard, the 20-foot deep pit. Then they are asked to carry a ball into the pit room and drop it on a target at the very bottom of the pit.

The reaction elicited by this VE could be due to vertigo, fear, or some other innate or learned responses. The reactions are well known in the literature and manifest as



**Figure 1. Side view of the virtual environment.**

**Figure 2. Subject wearing HMD and physiological monitoring equipment in the “Pit Room”.**

increased heart rate and skin conductance and decreased skin temperature. We hypothesize that the more presence users feel in the stressful virtual scene, the more physiological reaction they will exhibit. What causes this higher presence and higher physiological reaction? Is it due to a more realistic flow of visual information? Is it due to more coherence between the visual and haptic information? Is it due to the improved visual realism? We cannot answer these questions definitively. We have empirically shown, however, that physiological reaction and reported presence are both higher when we make improvements to the VE. For example, when we add a real wooden ledge that visitors can feel with their toes, we see a significant increase in heart rate. With the opportunity afforded to us by SIGGRAPH, attendees can find out for themselves: “How real can a virtual environment be?” and “How does this effect my physiological reactions?”

## References

- M. Meehan (2001). Physiological reaction as an objective measure of presence in virtual environments. Doctoral Dissertation. Computer Science. University of North Carolina, Chapel Hill, NC, USA.
- M. Meehan, B. Insko, M. Whitton, F. P. Brooks (2002). Physiological Measures of Presence in Stressful Virtual Environments. In Proc. of ACM SIGGRAPH 2002. ACM Press/ ACM SIGGRAPH.
- Hodges, L., B. Rothbaum, R. Kooper, D. Opdyke, J. Willford, T. Meyer and M. North (1994). Presence as the defining factor in a VR application. Technical Report GUV-94-06. Georgia Tech University, Graphics, Visualization, and Usability Center.

# Public Anemone: An Organic Robot Creature

Cynthia Breazeal

Andrew Brooks

Matt Hancher

Cory Kidd

John McBean

Dan Stiehl

Josh Strickon

MIT Media Lab, 77 Massachusetts Ave. NE18-5fl. Cambridge, MA 02139  
(617) 452-5601

cynthiab,zoz,mdh,coryk,jmcbean,wdstiehl,strickon@media.mit.edu

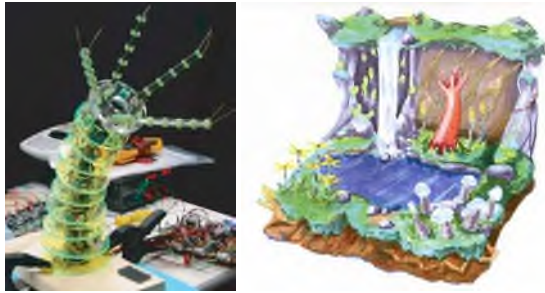


Figure 1: **Public Anemone, our interactive robotic character.** To the left is the underlying robot mechanism. To the right is conceptual drawing of the interactive terrarium installation.

## Abstract

We have created an articulated robotic creature and situated it within an interactive terrarium to explore the aesthetic, expressive, and interactive qualities that give robots an organic and engaging presence to people.

**Keywords:** autonomous robotics, interactive characters, organic robotics

## 1 The Motivation and Concept

Newly emerging robotic applications in the areas of entertainment, education, domestic assistants, etc. motivate the development of autonomous robots that can interact with people. Autonomous robots today (such as Sony’s robot dog Aibo or Honda’s humanoid robot Asimo) are hard-shelled, minimally expressive, and extremely limited in their ability to directly engage people. However, for people to welcome such robots into their daily lives, future robots must not only be proficient but also rewarding with which to interact. This project explores those aesthetic, behavioral, and interactive qualities that give a robot a life-like presence to people. Particular emphasis is placed on giving the robot organic qualities and making it highly expressive as advocated by Thomas and Johnson’s *The Illusion of Life*.

## 2 The Installation

To pursue these questions, we have taken primitive life as a metaphor. As such, we are building a robotic creature that is an intriguing blend of plant and animal. The robot is covered in a soft silicone skin and has a highly flexible body with sensate tentacles (see figure 1). The robot is situated in a 4 ft. by 5 ft. interactive terrarium that simulates the creature’s “alien home world” with a day-to-night cycle of ten minutes. During the day while the robot

creature is awake, the robot behaves as an autonomous creature, playfully engaging people about the exhibit as they try to interact with it as one might a terrarium pet. Using color cameras mounted within the terrarium, it visually perceives people via an assortment of features (motion, skin tone, color, depth, etc.). It responds to people’s gestures by orienting towards them, and smoothly tracking their movement. Particular care is paid to responding in a highly expressive manner and exhibiting a range of emotive responses such as a fear response if people are creating too much commotion about the creature. During the night cycle, the creature sleeps while other smaller, nocturnal creatures of the alien world awaken. Visitors interact with these simple robotic creatures through touch and proximity, causing them to respond through sound, light, and movement. These “nocturnal” interactions are managed by the show controls system, which models the terrarium as an intelligent stage.

## 3 Technological Innovations

Creating an interactive robot with organic qualities poses some interesting artistic and technological challenges. One critical aspect of giving a robot a life-like presence is designing it to move in a graceful, highly flexible, and articulated manner without sacrificing control accuracy and precision. The robot must move in both an expressive as well as a target-directed manner, so we are combining traditional robotic feedback control techniques with tools for scripting animations (e.g., telemetry for motion capture or animating a 3D graphical model of the robot) to ground hand-crafted animations with real-world physics. A visual system allows the robot to perceive people near the terrarium, and sensate tentacles allow it to feel its way about its environment. Based on our previous work with building sociable humanoid robots [Breazeal 2002], and building upon the code base of the Synthetic Characters Group [Burke et al. 2001] for interactive characters, these elements are tied together with a software architecture that integrates real-time perception, motivation, task-based behavior, and expressive movement.

## 4 The Team

The Robotics Presence Group at the MIT Media Lab is composed of specialists in artificial intelligence, sociable humanoid robotics, mechanical design, electrical engineering, machine perception, animatronic special effects, and human computer interaction.

## References

- BREAZEAL, C. 2002. *Designing Sociable Robots*. The MIT Press.
- BURKE, R., ISLA, D., DOWNIE, M., IVANOV, Y., AND BLUMBERG, B. 2001. Creature smarts: The art and architecture of a virtual brain. In *Proceedings of the Game Developers Conference*, 147–166.

# Regeneration of Real Objects in the Real World

Hiroto Matsuoka\*

Akira Onozawa

Hidegori Sato

Hisao Nojima

NTT Lifestyle and Environmental Technology Laboratories

## 1 Introduction

The vision of this work is to regenerate real objects that had existed at some moment in the past and/or at some remote location as if they have been transferred to the present across space and time. The objects could be museum pieces or items in stores, for example. For this work, we have developed a quick and fully automated system that can capture a three-dimensional image of real objects. This success brought us close to our goal of “regeneration of real objects in the real world.”

## 2 Technological Innovations

The system we have developed provides a scene in real time, where transferred real objects can exist naturally in the present real space. The system consists of functions for capturing, registration & sensing, and regeneration. The real objects are first captured by using our original 3D capturing system [Fig. 1]. A special feature of the system is its ability to capture 3D CG data from various real objects, such as plastic and porcelain items strong inter-reflection, translucent glass and acrylic resin objects, and so on. Before regeneration of the captured 3D CG data as virtual objects in real space, a sensor called the “sensor cube” measures environmental information about the location of the virtual objects. We have developed two kinds of sensor cubes, a geometry sensor and a light sensor [Fig. 2]. A viewer [Fig. 3] works to regenerate real objects in the real world by rendering them in the space at the “sensor cube” with ARToolKit [Schmalstieg et al. 2001] provided by the University of Washington. Since views of the virtual objects are generated using the sensor information, harmonization between the real world and virtual objects is achieved.

## 3 Demonstration

The installation we hope to introduce at SIGGRAPH2002 features “Noh” as the motif. “Noh” is a form of traditional Japanese drama in which the actors wear various kinds of mysterious masks. It is well-known that a single Noh mask can express various emotions depending on the viewing angle and the lighting. We captured

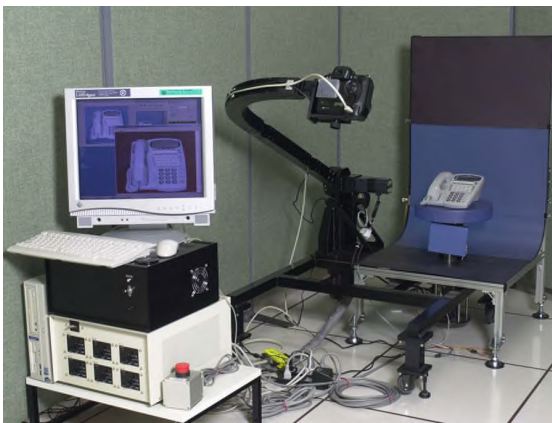


Fig. 1: The original 3D capturing system

images of Noh masks from various angles and in various lighting conditions over the period of the carving process. The captured images will be regenerated in the presence of the SIGGRAPH audience in real time anywhere the sensor cube is located as naturally as possible by sensing the light in the environment. In the exhibition, the carving process unfolds as distance between the audience and the cube changes. That is, a completed mask gradually reverts to a chunk of wood as the distance increases, and vice versa [Fig. 4]. We also hope to demonstrate the capturing of real objects using our system, and how the captured data are shown in other spaces using our viewer.

## 4 Conclusions

“Regeneration of real objects in the real world” is our ultimate goal. We want to “transfer” real objects into our presence beyond space and time and to let us interact with them in real time.

The proposed system can make objects presented in different places and different times appear to be ‘real’ and in the real space. The proposed system consolidates an extremely realistic three-dimensional CG, the reflection of environmental information of the actual space, and advanced Augmented Reality technology. This system, which can be applied and developed in versatile ways and means, will open a new dimension to our senses.

## References

SCHMALSTIEG, D., BILLINGHURST, M., AZUMA, R., HOLLERER, T., KATO, H., AND POUPYREV, I. 2001. “Augmented Reality: The Interface is Everywhere.” *Course Notes of SIGGRAPH 2001*



Fig. 2: The sensor cubes.



Fig. 3: The viewer.

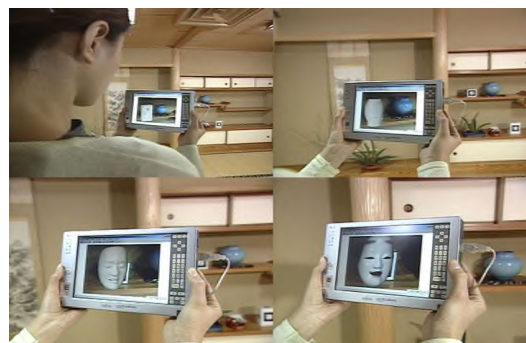


Fig. 4: The demonstration to show the mapping between the carving process of the mask and the distance between the observer and the sensor cube.

\*e-mail: hmatsu@aecl.ntt.co.jp



# SmartFinger : Nail-Mounted Tactile Display

Hideyuki ANDO

Japan Science and Technology  
Corporation (JST)  
smartfinger@star.t.u-tokyo.ac.jp

Takeshi MIKI

The University of Tokyo

Masahiko INAMI

The University of Tokyo

Taro MAEDA

The University of Tokyo / JST

## Abstract

"Smart Finger" is a novel type of tactile display for Augmented Reality (AR) that is wearable, like a press-on fingernail. This device allows the user to feel various textures while tracing his or her fingers along smooth objects. This wearable AR interface can supplement bump mapping information to real objects.

**Categories:** Emerging Technologies

**Keywords:** Tactile display, Bump mapping, Wearable device, SmartTool, Parasitic Humanoid

## 1 Introduction

The "SmartFinger" is a new type of tactile display for AR which is worn on the nail side of the finger. It does not inhibit our tactile sensation, since the ball of the finger is naked and we can feel the environment directly. The absence of any material between the object and the finger is a critical aspect of this design.

The "SmartFinger" can provide extra tactile sensations. For example, someone wearing this small device on his or her nail can touch a drawing on a flat sheet of paper and feel a virtual edge along the drawing's line, and virtual roughness from its texture pattern. This is a wearable device for AR, which adds bump mapping information to real surface objects using existing printing techniques.

## 2 Exposition

We propose a small, wearable tactile display composed of three components: (1) photodetectors to measure the changing environment beneath the finger in real-time, (2) a sensor to measure change in the fingernail's color from the nail side and detect the presence of external force, and (3) a "voice coil" to generate vibration on the nail, in order to realize various tactile sensations corresponding to these two types of sensory information.

## 3 Future Work

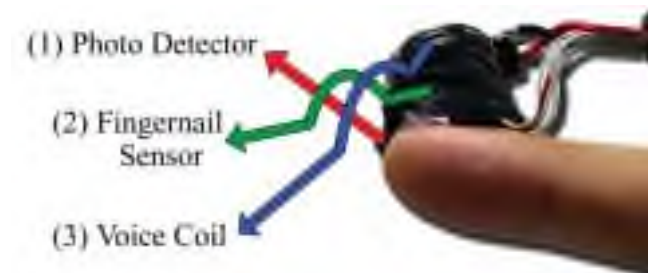
This research is part of The Parasitic Humanoid (PH) project to develop wearable robots that model nonverbal human behavior. A user of PH should be able to obtain data and analyze the results without suffering any inconvenience in daily life. To realize such a comfortably wearable PH, we need to develop measurement devices that don't interfere with human behavior. The SmartFinger fits this design criterion.

Of course, it can also be used as a hands-free pointing device with feedback. For example, since visually impaired people make extensive use of their tactile sense, they can use the device as a touch panel to detect information from a computer display.

**Acknowledgement:** This work is partially funded by a research grant from PRESTO, Japan Science and Technology Corporation (JST).



The "SmartFinger"



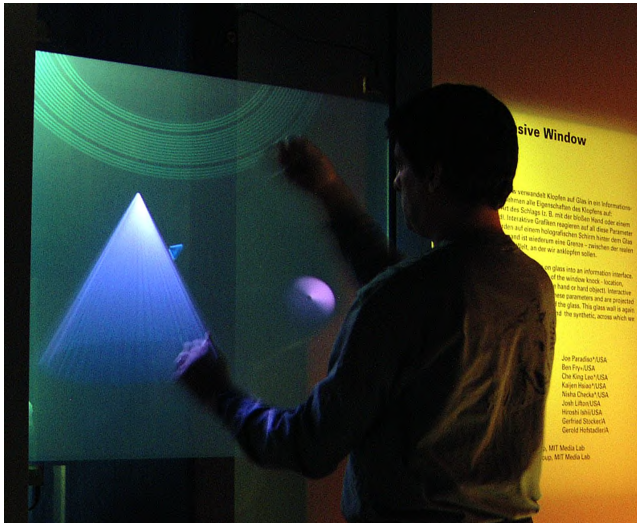
The three components of the "SmartFinger"

## References

- T. Nojima, M. Inami, Y. Kawabuchi, T. Maeda, K. Mabuchi and S. Tachi, "An Interface for Touching the Interface", ACM SIGGRAPH 2001 Conference Abstracts and Applications, p.125, 2001
- Mascaro S. and Asada H., 2001, "Photoplethysmograph Fingernail Sensors for Measuring Finger Forces Without Haptic Obstruction," IEEE Transactions on Robotics and Automation, Vol. 17, No. 5, pp. 698-708
- T. Maeda, H. Ando, J. Watanabe, Y. Nonura and T. Miki, "A behavior modeling with wearable robotics "The study of Parasitic Humanoid (VI)", The 6th VRSJ Annual Conference, pp. 153-154 (Japanese)

# The Interactive Window

Joseph A. Paradiso<sup>\*</sup>, Che King Leo<sup>\*</sup>, Nicholas Yu<sup>\*</sup>, Marc Downie<sup>+</sup>  
Responsive Environments Group<sup>\*</sup>, Synthetic Characters Group<sup>+</sup>  
MIT Media Laboratory  
{joep, cheking, nhyu, marcd}@media.mit.edu



We have developed a very simple retrofit to a large display surface that enables knocks or taps to be located and characterized (e.g., determining type of hit – metallic tap, knuckle tap, or bash – and intensity) in real time. We do this by analyzing the waveforms captured by 4 piezoelectric transducers (one mounted in each corner of the surface) and a dynamic microphone (mounted anywhere on the glass) in a digital signal processor. Differential timing yields the position, frequency content infers the kind of hit, and peak amplitude reflects the intensity. This technique was first explored in collaboration between Paradiso and Ishii [Ishii et al. 1999] to make an interactive ping-pong table. Moving to glass display surfaces introduced significant problems, however – knuckle taps are low-frequency impulses that vary considerably hit-to-hit, and the bending waves propagating through the glass are highly dispersive. A heuristically-guided cross-correlation algorithm [Paradiso et al. 2002] was developed to counteract these effects and provide spatial measurements that can resolve knuckle impacts to within  $\sigma = 2-4$  cm (depending on the material thickness) across a 2-meter sheet of glass. As the requisite hardware is minimal, and everything is mounted on the inside sheet of glass, this is a very simple retrofit to, for example, store window displays, ushering in an entirely new concept of interactive window browsing, where passers-by can interact with information on the store's products by simply knocking. We have explored this concept in retail, where one of our trackers was installed on the main display window of an American Greetings store near Rockefeller Center in Manhattan for this year's Christmas-Valentine's Day season (right figure), and in museums (e.g., left figure, which shows the system running at the Ars Electronica Center in Linz, Austria).

We plan to augment the contact interaction of the knocking system with a noncontact system that detects the presence and activity of participants in front of the screen. Although we have used motion radar systems before for measuring user activity (e.g., [Paradiso et al. 1997]), we are now developing a small, low-

power ranging radar system that can determine the distance to users as they approach the screen. Unlike IR, vision or sonar systems, this technique isn't sensitive to lighting, clothing, or clutter. It can also sense through opaque, nonconductive material, such as plastic, wood, or wallboard. We plan to place 2-3 radar antennae behind the frosted glass display and outside of the light cone of the projector. This will allow the system to zone the users as they approach, adjusting the audiovisual content accordingly.

The participants' activity will result in accompanying music and graphics. Free gesture will produce gentle sounds and amorphous cloud-like graphics that become more frenetic with decreasing range. Contact interaction will materialize discrete audiovisual events. In particular, a non-representational, non-photorealistic rendering style will be used that constructs solid waves of color out of lines and points, in which shapes and details repeatedly condense and evaporate. This visualization, which builds its images out of overlapping perturbations of textures, is particularly appropriate for an installation where the primary interaction is the act of striking the display surface.

## References

- ISHII, H., ET AL., 1999. PingPongPlus: Design of an Athletic-Tangible Interface for Computer-Supported Cooperative Play. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI '99)*. ACM Press, 1999. 394-401.
- PARADISO, J., ET AL., 2002. Passive Acoustic Sensing for Tracking Knocks Atop Large Interactive Displays. To appear in the *Proceedings of the 2002 IEEE Sensors Conference*, June 2002.
- PARADISO, J., ET AL., 1997. The Magic Carpet: Physical Sensing for Immersive Environments. In the *Proceedings of the Conference on Human Factors in Computing Systems (CHI '97), Extended Abstracts*. ACM Press, 1997. 277-278.

# The Virtual Showcase: A Projection-Based Multi-User Augmented Reality Display

**Oliver Bimber\***  
Fraunhofer Center for  
Research in Computer  
Graphics,  
Providence, RI, USA,  
obimber@crcg.edu

**Bernd Fröhlich**  
Bauhaus University Weimar,  
Germany  
bernd.froehlich@medien.  
uni-weimar.de

**Dieter Schmalstieg**  
Vienna University of  
Technology, Austria,  
dieter@cg.tuwien.ac.at

**L. Miguel Encarnação**  
Fraunhofer Center for  
Research in Computer  
Graphics,  
Providence, RI, USA,  
me@crcg.edu

\*responsible author

The Virtual Showcase is a new projection-based and application-specific Augmented Reality display that offers an innovative way of accessing, presenting, and interacting with scientific and cultural content. Conceptually, the Virtual Showcase is compatible with the conventional showcases used, for instance, by museums. However, it allows the display of computer generated 3D graphics and animations together with real artifacts within the same space. From the technological point of view, the Virtual Showcase provides perspective correct stereoscopic viewing for multiple users, high resolution, low parallax (reflected projection plane inside the showcase), and support for mutual occlusion between real and virtual objects.

The Virtual Showcase consists of a convex assembly of half-silvered mirrors and a graphics display. The showcases' contents are illuminated with a controllable light source while view-dependent stereoscopic graphics are presented to head-tracked observers. Using a regular light bulb to illuminate physical objects located inside the showcase results in a semi-transparent appearance of the virtual objects. The real objects can be seen through the overlaid graphics. This results in conflicting occlusion cues. The use of video projectors as illumination sources allows the generation of shadows on the physical objects surface where the corresponding graphical content is overlaid. Appearance parameters of the projected light such as color, blending, and geometry can be controlled statically or dynamically (considering the current point of view during rendering) on a pixel-precise basis. This results in a realistic mutual occlusion of real objects by virtual ones. We use a digital light projector (DLP) and additional full mirror optics to realize a pixel-precise illumination. The full mirror optics are used to redirect the projected light into the Virtual Showcase.

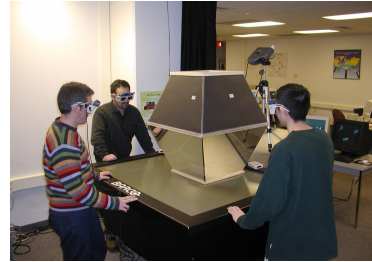
The Virtual Showcase turns the exploration of cultural and scientific objects into an interactive process, which significantly enhances the visitor's museum experience and facilitates the learning process. Visitors can change from their passive role of an observer to the active role of an investigator. Multiple users can interact simultaneously with virtual content displayed in a single showcase stimulating communication and encouraging joint discovery of the presented pieces.

## Reference

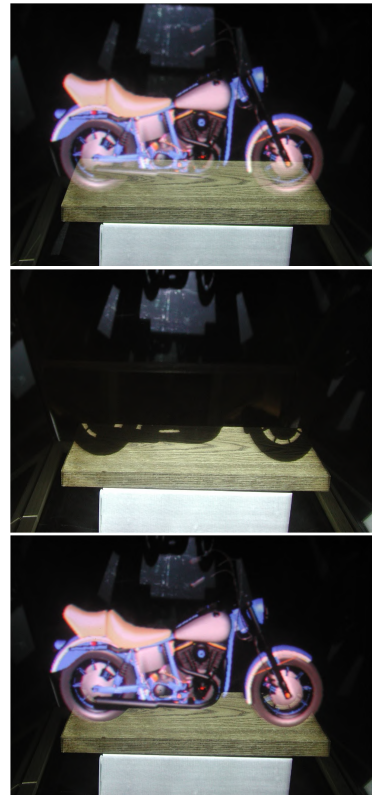
Bimber, O., Fröhlich, B., Schmalstieg, D., and Encarnação, L.M. "The Virtual Showcase." *IEEE Computer Graphics & Applications*, vol. 21, no.6, pp. 48-55, 2001.

## Acknowledgements

The project is supported by the European Union (IST-2001-28610) and the U.S. Army (BAA NAWCTSD 97-02).



The Virtual Showcase prototype supports pixel-precise illumination..



Illumination through a simple light bulb causes conflicting occlusion cues (top) and virtual objects appear always semi-transparent. To overcome this, a view-dependent shadow is generated on the physical objects' surface (center) before the virtual object is overlaid (bottom).



# TWISTER: A Media Booth

Kenji Tanaka Junya Hayashi Yutaka Kunita Masahiko Inami Taro Maeda Susumu Tachi  
 Graduate School of Information Science and Technology, The University of Tokyo  
 e-mail: twister@star.t.u-tokyo.ac.jp  
 http://www.star.t.u-tokyo.ac.jp/projects/mutel/

TWISTER (Telexistence Wide-angle Immersive STEReoscope) is an immersive full-color autostereoscopic display, designed for a face-to-face telecommunication system called “mutual telexistence”, where people in distant locations can communicate as if they were in the same virtual three dimensional space.

The basic idea of mutual telexistence is the projection of human beings into a virtual environment in real time[Tachi et al. 1996][Kunita et al. 1999]. Each human user resides inside a booth, which displays a live full-color panoramic and autostereoscopic images to the user, and at the same time, captures images of the user from every angle. Autostereoscopic here means there are no obstacles to hide the user’s face, and therefore, the face of the observer can be clearly seen. With multiple booths, people at remote locations can meet together as if they were close at hand.

TWISTER has rotating display units surrounding the observer. One unit consists of two LED (Light-Emitting Diode) arrays and a parallax barrier. Each LED array consists of pairs of red, green, and blue LED’s, and displays time-varying patterns so that the observer can perceive an image. Due to the use of LED, TWISTER can be used as a display of high intensity full-color images for use in a bright room. The rotation of the display unit makes it a wide angle display. In fact, the angle of view as a normal display is 360 degrees.

The key device for autostereopsis is the rotating parallax barrier. One of the LED arrays is for the left eye, and the other is for the right eye. Because the parallax barrier obscures the opposite side LED emission, different images are shown to the left and right eyes. The angle of view as a stereoscopic display depends on the direction and the position of the observer. If the head of the observer is fixed, it exceeds 120 degrees. Thus, the observer can experience panoramic and autostereoscopic images without bothersome eyewear.

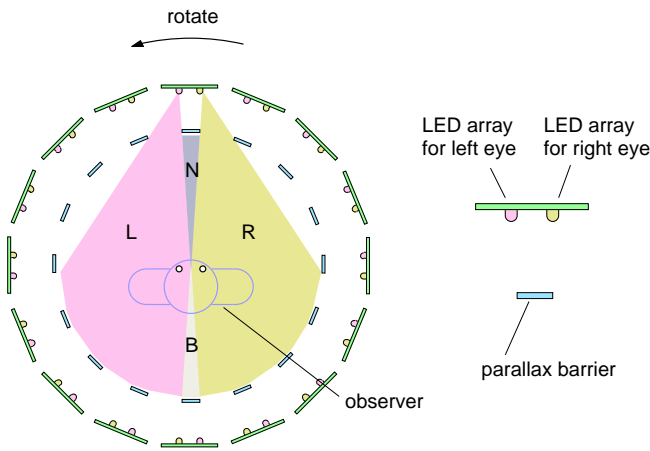


Figure 1: The principle of the rotating parallax barrier. In areas L and R, only images for the left and right eye can be seen, respectively. In area B, both images can be seen, while in area N no image can be seen.



Figure 2: A user stands inside a booth (left) and sees the figure of other person (right).

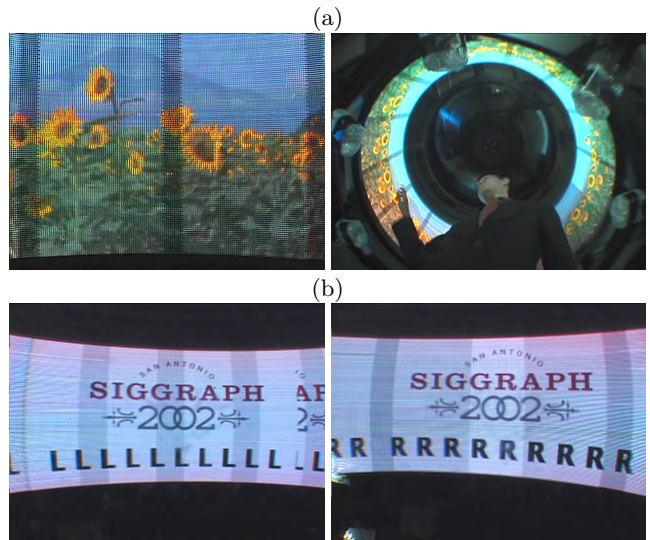


Figure 3: (a)The displayed full-color natural images. (b)TWISTER provides separate light sources to the left and right eyes.

*Acknowledgments* This research has been conducted as a part of the CREST Telexistence Communication Systems Project supported by JST (Japan Science and Technology Corporation)

## References

KUNITA, Y., INAMI, M., MAEDA, T., AND TACHI, S. 1999. Prototype system of mutual tele-existence. In *ACM SIGGRAPH '99 Conference Abstracts and Applications*, 267.

TACHI, S., MAEDA, T., YANAGIDA, Y., KOYANAGI, M., AND YOKOYAMA, H. 1996. A method of mutual tele-existence in a virtual environment. In *Proceedings of the ICAT'96*, 9–18.

## Ultrasound Visualization with the Sonic Flashlight

Damion Shelton  
Carnegie Mellon University  
dmskelto@andrew.cmu.edu

George Stetten  
University of Pittsburgh /  
Carnegie Mellon University  
george@stetten.com

Wilson Chang  
University of Pittsburgh  
wmcst6+@pitt.edu

From the discovery of X-rays over a century ago, clinicians have been presented with a wide assortment of imaging modalities yielding maps of localized structure and function within the patient. Some imaging modalities are *tomographic*, meaning that the data are localized into voxels, rather than projected along lines of sight as with conventional X-ray images. Tomographic modalities include magnetic resonance (MR), computerized tomography (CT), ultrasound, and others. Tomographic images, with their spatially distinct voxels, are essential to our present work.

New techniques to display tomographic images have lagged behind the development of imaging modalities themselves. In the practice of medicine, the standard method of viewing an image is still to examine a photographic film or an electronic screen rather than to look directly into the patient. The ability to view images at their actual locations within the patient could have a broad impact on the diagnosis and treatment of disease by providing *in situ* guidance for invasive procedures.

We have developed a new method of guidance that displays an image *in situ* within the patient in real time. The method, which we call *real time tomographic reflection* (RTTR), employs a half-silvered mirror to combine a direct view of the patient with a reflection from the display of an imaging device. The reflected image appears to be floating within the patient at its actual location, and the user perceives this 3D effect using natural, viewpoint independent, stereoscopic vision. Unlike previous attempts at such a visual merger, RTTR does not require tracking the patient or the viewer and involves no head-mounted apparatus. By presenting the patient, the operator's hands, the invasive tool, and the image in a single perceptual environment, the operator may take full advantage of natural hand-eye cooperation.

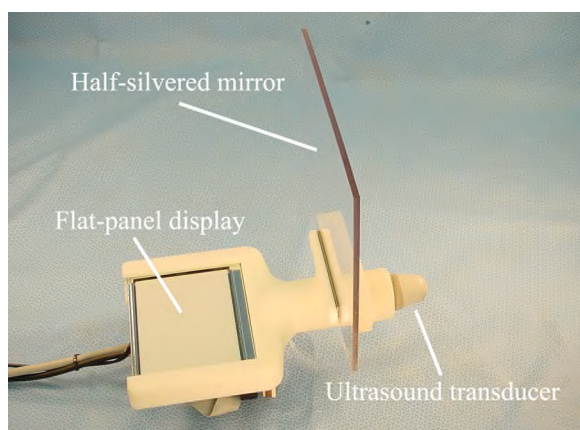


Figure 1 – Components of the Sonic Flashlight.

Our present implementation of RTTR uses ultrasound as its imaging modality and is named the *Sonic Flashlight* because the reflected ultrasound image appears to emanate from the transducer. Because of its small size, it can be used single-handed, leaving the other hand free to perform medical procedures such as ultrasound guided needle biopsy. Unlike many other tomographic visualization techniques, such as those involving head tracking,

third-party observers not directly manipulating the device also view the ultrasound in its correct physical location, and the unit may be passed between multiple personnel during a procedure without requiring recalibration or transfer of bulky external equipment.

Preliminary tests of the Sonic Flashlight have demonstrated the ability to hit targets using natural hand-eye coordination inside common ultrasound phantoms such as water-filled balloons and turkey breasts. Recently, we have collaborated with an ophthalmologist to test the ability of the Sonic Flashlight to guide a needle during a procedure known as retrobulbar injection, during which a long needle is used to inject a therapeutic substance behind the eyeball. Further clinical testing is planned for a wide variety of procedures, including catheter insertion and tumor biopsy.

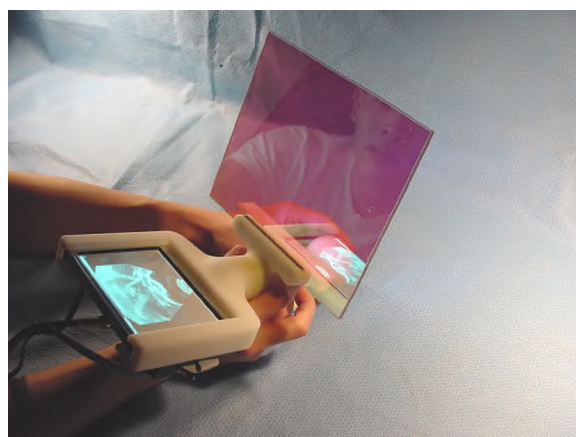


Figure 2 – The Sonic Flashlight in operation.

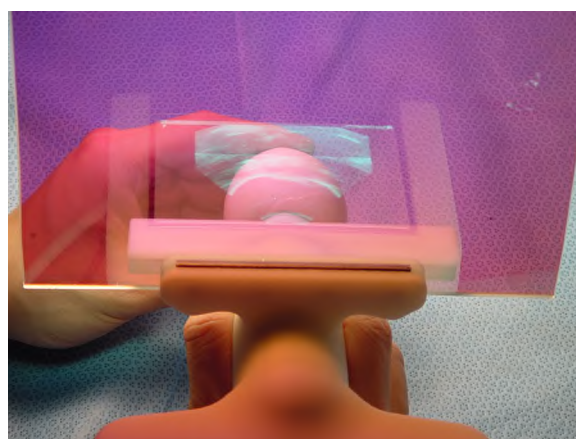


Figure 3 – An ultrasound scan of a water-filled balloon as viewed through the mirror.

# Virtual Chanbara

Daijiro Koga  
The University of Tokyo  
chanbara@star.t.u-tokyo.ac.jp

Takahiro Itagaki  
The University of Tokyo  
chanbara@star.t.u-tokyo.ac.jp

## Abstract

Chanbara is the name of our game, which reproduces the original ancient art of "samurai sword fighting". For this work, we created a new force feedback device called "GEKI2". The player holds the sword and wears an HMD. In Chanbara, the player walks around a virtual arena to battle against CG characters.

**Keywords:** Chanbara, force display, virtual reality

## 1 Introduction

Kendo, the art of Japanese swordsmanship, has a long and rich history.

Since old-fashioned training with real steel swords and hardwood swords caused so many unnecessary injuries and deaths, harmless bamboo practice swords were created around 1710 developed by Japanese armors and Japanese sword masters. As can be imagined, the original bamboo practice swords and protectors were quite primitive and of simple construction.

Now we developed a new type Kendo with virtual reality technology, which uses an HMD and a novel force display, instead of a head protector and a bamboo practice sword.

## 2 GEKI2

GEKI2 is made up of two motors that create two rotational moments when the player grasps it. Usually, these two motors are spinning, and to generate an impact, GEKI2 will stop both motors immediately. Through the conservation of angular momentum, the inertia will reach the player's hand, so the hand feels an impulsive force. Immediately after stopping, the motors begin spinning again, although they accelerate gradually, in order not to pose a noticeable force to the hands. We control the acceleration by PWM, making it possible to change the strength of impacts by changing the duration of braking time.

A virtual enemy, virtual sword and virtual arena are displayed with the HMD. The graphics are updated based on the information from the magnetic sensors.

The player attacks the enemy and blocks the enemy's attacks with the virtual sword. GEKI2 returns an impulsive force to the player's hands, synchronized with the attacks. We detect whether an attack hits by using the information from the magnetic sensors. The game finishes once the player conquers several enemies, or is defeated (about a few minutes). As we use HMD in this work, the audience can't see how the player is fighting, so we make a mixed graphics display of the player and computer graphics, shown here. Now, the audience can see how players are fighting and cheer then on. This exciting demonstration is enjoyable for people of all ages.

## 3 Future Work

We are planning to downsize this device. In addition, Virtual Chanbara is suitable for networked play. In the future, a Chanbara master can instruct worldwide children with this system, and an international match can be held on the Internet.

## Acknowledgements

This study was supported in part by a grant for the Inter-collegiate Virtual Reality Contest (IVRC) from Gifu Prefecture and Kakamigahara City.





# Panels – Introduction

The SIGGRAPH 2002 Panels program is a highly interdisciplinary program that sparks animated discussion and provides through-provoking insights from some of the top professionals in the interactive and graphics world. At SIGGRAPH 2002, we have twelve panels on topics that span the gamut of Computer Graphics and Interactive Techniques. The Panels program gives SIGGRAPH attendees the opportunity to listen and participate in discussions about the future of our field.

The panel “The Future of Computer Graphics: An Enabling Technology?” explores what should be next for Computer Graphics and Interactive Techniques. The question is this — Is Computer Graphics dead? Probably not, but there are some in our community who fear that Computer Graphics research is exploring smaller and smaller problems with corresponding smaller and smaller importance. The panel on “The Future of Computer Graphics” will discuss areas of Computer Graphics and Interactive Techniques that require the best research that the field can give.

As always, the advance of technology is shaping how we think and interact with digital imagery. Moore’s Law is still in effect, delivering faster computers and bigger memory in smaller packages every year. Faster computers imply new capabilities, and the panel “When Will Ray-Tracing Replace Rasterization” seeks to upset long-held assumptions about rendering on desktop machines. For mobile applications, the new trend towards improved graphics capabilities on small mobile displays is discussed in “Unsolved Problems in Mobile Computer Graphics and Interaction”. At the opposite end of the display scale, the panel “Graphics in the Large” explores issues related to interacting with large projection displays.

Another broad theme in the SIGGRAPH 2002 Panels program is animation. The panel “Digital Humans: What Role Will they Play” discusses the rise of digital humans that are difficult to distinguish from the real thing, at least in photos. What is next in the area of human animation? Closely related in terms of topic is the panel “Animation’s Turning Tide”, which will discuss the recent embrace of 3D animation by professional Hollywood animators. Continuing on the animation theme, the panel “How Will Motion Capture Affect Animation” discusses MoCap in relation to film production, and how it should improve to better serve the needs of animators. The fourth panel session on animation is a light-speed recap of the papers presented at the co-located ACM SIGGRAPH Symposium on Computer Animation.

Rounding out the program, we have panels that exhibit the wide range of disciplines and interests contributing to SIGGRAPH. “Extending Interface Practice: An Ecosystem Approach” discusses the concepts behind new interactive systems that bring new meaning to the user. “Interactive Stories: Real Systems, Three Solutions” discusses the storytelling issues in the context of systems built by the panelists. The panel “The Demo Scene” explores a heretofore underground movement of programmers and artists who build extraordinarily compact visual demos.

Finally, the SIGGRAPH 2002 Panels program wraps up with the provocative statement “Games: the Dominant Medium of the Future”. Will video games replace television and motion pictures as the primary entertainment medium in the 21st Century? SIGGRAPH 2002’s closing panel will grapple with that thought.

Christopher D. Shaw, Chair

# Panels – Committee

## Chair

Christopher Shaw  
Georgia Institute of Technology

## Panels Jury

John Buchanan  
Electronic Arts Canada

David S. Ebert  
Purdue University

Diane Gromala  
Georgia Institute of Technology

Lou Harrison  
North Carolina State University

Holly Rushmeier  
IBM T. J. Watson Research Center

# When Will Ray-Tracing Replace Rasterization?

Kurt Akeley\*  
Stanford University  
NVIDIA Corporation  
Panelist

David Kirk†  
NVIDIA Corporation  
Panelist

Larry Seiler‡  
ATI Research, Inc.  
Panelist

Philipp Slusallek§  
Saarland University  
Panelist

Brad Grantham¶  
Silicon Graphics Incorporated  
Moderator

## 1 Motivation

Ray-tracing produces images of stunning quality but is difficult to make interactive. Rasterization is fast but making realistic images with it requires splicing many different algorithms together. Both GPU and CPU hardware grow faster each year. Increased GPU performance facilitates new techniques for interactive realism, including high polygon counts, multipass rendering, and texture-intensive techniques such as bumpmapping and shadows. On the other hand, increased CPU performance and dedicated ray-tracing hardware push the potential framerate of ray-tracing ever higher.

Depth-buffering rasterization made scanline geometric hidden-line techniques obsolete because it was easy to implement and dropping memory prices made implementations affordable, even though Sutherland, et al., concluded it was hopelessly inefficient.

## 2 Question for the Panel

Will the simplicity and/or increasing performance of ray-tracing make rasterization obsolete, and when? Speakers will address the future of rasterization versus ray-tracing/ray-casting techniques based on their broad and diverse industry experience and individual viewpoints as leaders in the graphics community.

## 3 Kurt Akeley

Rasterization hardware and the associated graphics standards, such as Direct 3D and OpenGL, will be around for a long time. Ray tracing will become feasible, outside of niche markets serviced with custom hardware, only as it can be implemented using rasterization hardware infrastructure. Because rasterization hardware is advancing in performance, flexibility, and programmability so rapidly, we can expect dramatic improvements in the image quality of interactive systems during the next few years. Applications will introduce ray tracing algorithms gradually, synthesizing rasterization, ray tracing, and other global shading techniques to obtain the best overall results.

Kurt Akeley works part time at NVIDIA Corporation, where he is a member of the graphics architecture team. He spends the rest of his week at Stanford, working toward the completion of the electrical engineering Ph.D. that he put on hold in 1982 to co-found Silicon Graphics.

During his 19 years at Silicon Graphics Kurt lead the development of several high-end graphics systems, including GTX, VGX,

\*e-mail: kurt\_akeley@acm.org

†e-mail: davidk@nvidia.com

‡e-mail: lseiler@ati.com

§e-mail: slusallek@cs.uni-sb.de

¶e-mail: grantham@sgi.com

and RealityEngine. He also lead the development of OpenGL, and he continues to be involved with OpenGL's evolution. Kurt's past SIGGRAPH participation includes course presentations, panel presentations, and participation in the technical program, which he chaired in 2000. Kurt is named on thirteen patents, is a fellow of the ACM, and in 1995 was the recipient of the ACM SIGGRAPH Computer Graphics Achievement Award.

## 4 David Kirk

Ray Tracing vs. Rasterization is an interesting question, although perhaps a bit ill-posed. Ray tracing is used both as a visibility technique and a catch-all to describe a variety of global illumination techniques. Usually these techniques are CPU-based. Rasterization, typically depth buffering, is primarily used as a visibility technique, although some texture-based shadowing, illumination, and reflection techniques can make use of rasterization as the workhorse. Usually these techniques are GPU-based.

I'll be interested in discussing a bigger question, though: "When will hardware graphics pipelines become sufficiently programmable to efficiently implement ray tracing and other global illumination techniques?". I believe that the answer is now, and more so from now on! As GPUs become increasingly programmable, the variety of algorithms that can be mapped onto the computing substrate of a GPU becomes ever broader. As part of this quest, I routinely ask artists and programmers at movie and special effects studios what features and flexibility they will need to do their rendering on GPUs, and they say that they could never render on hardware! What do they use now: crayons? Actually, they use hardware now, in the form of programmable general-purpose CPUs. I believe that the future convergence of realistic and real-time rendering lies in highly programmable special-purpose GPUs.

David is Chief Scientist and Vice President of Architecture at NVIDIA. He was previously Chief Scientist and head of technology, developing video game software for Crystal Dynamics, and prior to that worked on developing graphics hardware for engineering workstations at Apollo/Hewlett-Packard. David holds B.S. and M.S. degrees from MIT and M.S. and Ph.D. degrees from the California Institute of Technology, and is the author/inventor of over 100 technical publications and patents in the area of computer graphics and hardware.

## 5 Larry Seiler

Ray-tracing will never replace rasterization. Both will be subsumed by something that is better than either, for three reasons. First, the two algorithms do not solve the same problem. Ray-tracing excels at modeling a physical environment. Rasterization excels at splicing many different algorithms together (that's a strength, not a weakness) to achieve a desired result that may not match physical reality. Second, memory bandwidth doesn't increase as fast as CPU performance. Rasterization algorithms are better able to produce the reference locality and good cache behavior that is required to use memory bandwidth efficiently. Finally, rasterization is typically better at tolerating long latencies, which are getting worse as GPUs and graphics models get more complicated.

The Next Big Thing needs to have the strengths of both algorithms and needs to solve a problem common to both: aliasing! Standard rasterization and standard ray-tracing algorithms create point samples that must be post-filtered, instead of correctly pre-filtering the model. They also share the inefficiency of repeating similar operations many times when rendering multiple frames. Possible futures are suggested by work on beam-casting, IBR, and by the Shadermaps algorithm, which saves intermediate computations at multiple resolution levels for use in multiple frames. These and other techniques should be extended to allow mixing ray-traced, rasterized, and even volumetric techniques, probably implemented in ways that blur the distinction between ray-tracing and rasterization.

Larry Seiler is a hardware architect at ATI Research, Inc. Previously, Larry worked at Real Time Visualization developing volumetric ray tracing algorithms. Larry described this work at the 1999 MicroProcessor Forum and is a co-author of "The VolumePro Real-Time Ray-casting System" (SIGGRAPH 1999). Other paper credits include "Quadratic Interpolation for Near-Phong Quality Shading" (SIGGRAPH 1998 technical sketch) and co-authorship of "Neon: A Single-Chip 3D Workstation Graphics Accelerator" (1998 EUROGRAPHICS / SIGGRAPH Workshop on Graphics Hardware, best paper). Larry holds B.S. and M.S. degrees from Caltech and a Ph.D. degree from MIT. Larry is an inventor for 19 issued patents.

## 6 Philipp Slusallek

The ray-tracing algorithm does have a number of key benefits that make it significantly more attractive to use than the rasterization algorithm of today's graphics hardware:

- **Flexibility:** Ray-tracing can accurately compute exactly those rays required for some effect, while rasterization is highly limited in what it can compute efficiently.
- **Generality:** Because of this flexibility, ray-tracing is the perfect match for many advanced rendering algorithms and effects required for future 3D graphics.
- **Approximations:** Rasterization has to rely on approximations and special hacks for almost every rendering effect. These hacks are usually non-intuitive, non-trivial to find and implement, and are costly or even impossible to combine with other approximations.
- **Plug & Play:** Ray-tracing directly uses high-level shaders (such as RenderMan) that can be combined with arbitrary complexity. Even large scenes with many, complex shaders simply work as expected without further work or tweaking of parameters.

- **Scalability:** Ray-tracing scales extremely well both in scene size (logarithmic versus linear for rasterization) and in terms of the number of processing elements it can make use of. The latter issue will be important for future hardware where billions of transistors in a chip must be kept busy.

Due to these arguments I am convinced that in the future, 3D graphics will be based on ray-tracing. The biggest issue along this road is the question about the best API and the transition to it.

What remains is the question about hardware: Will ray-tracing and 3D graphics run on the GPU, CPU, or on some new special-purpose hardware? The stream-processing model of the GPU has many arguments in its favor and will most likely support some form of ray-tracing in the near future. However the CPU guys are not standing still either and will likely support high-performance ray-tracing on single CPUs in the not too distant future, too.

It will be interesting to watch the two groups, each striving to build the best hardware support for ray-tracing and the future of 3D graphics.

Philipp is full professor at the computer graphics lab of Saarland University. From 1998 to 1999 he was visiting assistant professor at the Stanford University graphics lab. He received a Diploma in physics from the University of Tübingen, Germany in 1990 and a PhD in computer science from the University of Erlangen, Germany in 1995. He was responsible for the design of a commercial 2D/3D CAD package and has been leading the Vision project, a large object-oriented and physically-based rendering system. His current research activities focus on interactive ray-tracing and interactive lighting simulation. Other research topics include design of a network-integrated multi-media infrastructure, consistent illumination in virtual environments, physically-based and realistic image synthesis, and object-oriented software design. Philipp organized the SIGGRAPH 2001 course "Interactive Ray-Tracing".

## 7 Brad Grantham

Rasterization has always represented speed because of its direct forward mapping from triangles to screen pixels. Over its many years of dominance in interactive rendering, lots of techniques have been incorporated for increased realism. Ray-tracing, however, provides global illumination and high-quality surface appearance at a much lower price in implementation. Ray-tracing will continue to increase in speed as Moore's law carries us forward and as commodity ray-tracing hardware appears, but rasterization relies on the creativity and cleverness of engineers and artists in the face of exponentially increasing complexity. Perhaps some new algorithm will triumph over both in the long term but my prediction consists of the adoption within five to ten years of ray-tracing over rasterization (with perhaps an intermediate hybrid) through a dedicated device much like today's AGP 4x graphics adapter and an open standard API embodying both ease-of-use and performance much as OpenGL has provided for rasterization.

Brad Grantham works in the Applications Engineering division at SGI, specializing in helping application developers use OpenGL and SGI's toolkits for visualization. Previously, Brad worked for VA Linux Systems on graphics workstation development and before that in the Advanced Graphics Software division of Silicon Graphics, Inc. on scene graphs and toolkits.

Brad spoke at the "Advanced Rendering Techniques in OpenGL" course at SIGGRAPH from 1997 through 2000 and organized the "OpenGL on Linux" course at the Fall 2000 Game Developers' Conference. Brad also designed and implemented the open source "ACTC" API and library for creating triangle strips from triangles.

# Digital Humans: What Roles Will They Play?

Norman I. Badler  
University of  
Pennsylvania  
Panelist

Nadia Magenat-  
Thalmann  
MIRALab, University  
of Geneva  
Panelist

Laurie McCulloch  
Digital Animation  
Group  
Panelist

Evan Marc Hirsch  
Electronic Arts  
Panelist

Phil LoPiccolo  
Computer Graphics  
World  
Moderator

## 1. Introduction

Computer graphics technology has progressed to the point where it is possible to create digital humans that are virtually indistinguishable from the real items. The potential benefits are immense, but there are implications to consider as well, in a range of applications that includes film, video, the Web, and gaming. This panel of experts from diverse disciplines of computer graphics will discuss how far we have come in the use of digital humans, where they are heading, and what they will mean to us.

### 1.1. Questions

What are the positive ramifications of digital humans in your field?

What are some of the controversial aspects of the technology?

Which applications or examples represent state-of-the-art digital humans in your area of computer graphics? Where is state of the art headed?

How can common technologies in the various fields in which digital humans are used be applied for different purposes? Put another way, how might technology transfer occur?

## 2. Norman I. Badler, panelist

What will virtual humans be like in 100 years? Will there be any difference? No. This may be the last generation that sees a difference between real and virtual things. Will digital humans decide whether to do our bidding or will they be our new slaves? Will they demand emancipation? Will they be an outlet for our baser drives or desensitize us to them (games and violence; sex and pornography—cause-and-effect relationships?). Do they accommodate to us or do we accommodate to them? Will they need beliefs and axioms (dogma)? Will we be their gods?

### 2.1 Badler bio

Norman I. Badler is a Professor of Computer and Information Science at the University of Pennsylvania, and is Associate Dean for Academic Affairs in the School of Engineering and Applied Science. Active in computer graphics since 1968 with more than 200 technical papers, his research focuses on human figure modeling, manipulation, and animation control in real-time 3D graphics. Badler's current research interests include animation via simulation, embodied agent software, human-computer interfaces, and computational connections between language and action. Badler received a Ph.D. in Computer Science in 1975 from the University of Toronto. He is Co-Editor of the journal *Graphical Models*. Badler also directs the Center for Human Modeling and Simulation. Among the Center's achievements are the human modeling software system "Jack" that was the basis for a spin-off company in 1996 (Transom Technologies, bought by Engineering Animation, and now a part of EDS). Badler has been active in SIGGRAPH since 1976, being Conferences Tutorial Chair in the late 1970s and national Vice Chair from 1978 to 1981.

## 3. Nadia Magnenat-Thalmann, panelist

Since the time we simulated Marilyn Monroe and Humphrey Bogart in the well-known film "Rendez-vous in Montreal" in 1987, featuring for the first time two legendary stars, more and more virtual humans have been created that showed a more realistic appearance. They also have more natural functions, such as showing emotions, walking, talking, etc. What is the next step? First of all, virtual humans should not be made mainly by hand, i.e. with a lot of real human designer interaction. They should be modeled more automatically and above all possess knowledge themselves. They should know how to behave in any circumstance as we real humans do. They should also retain memory, which means knowing who is interacting with them, and then behave accordingly. They should dress themselves, comb their hair, decide on sports activities, and have interactions between each other automatically. This will allow the creation of a virtual world where virtual people are conscious of their environment and can behave, act and react in the proper way without a huge amount of manual work necessary on the part of humans. The applications are endless: We will be able to simulate any situation before it is created for real, such as in manufacturing or cultural heritage. Digital humans will interact with us on all topics and explain to us, with real-time simulations, what could happen in a given event. Also, an area where they will take an increasing role is the medical field. We will simulate internal individual organs with biomechanical models and we will soon be able to be operated on, on our own virtual reconstructed human model before any surgery is performed on our real body. When real surgery is performed, it will be done mainly with the help of 3D real-time organ reconstruction and deformation. This will allow a more precise knowledge of individual bodies and provide better results and a chance to live longer and in better shape. In the long run, real humans will have more and more prosthetic devices—chips to increase the duration and quality of life. On their side, the virtual humans, functioning with chips, will move slowly to biological hardware. The difference in nature between the virtual and real humans will decrease slowly. Fantasy or reality? Only the future can provide the answer.

### 3.1 Magnenat-Thalmann bio

Professor Nadia Magnenat-Thalmann has pioneered research into virtual humans over the last 20 years, participating in and demonstrating some of the most spectacular state-of-the-art developments in the field, and is responsible for the rigorous and intensive academic research programs that made them possible. Nadia Magenat Thalmann studied at the University of Geneva and obtained several bachelor degrees, including Psychology, Biology, and Computer Science. She received a MS in Biochemistry in 1972 and a PhD in Quantum Physics in 1977. From 1977 to 1989, she was a Professor at the University of Montreal in Canada. In 1989, she founded MIRALab, an interdisciplinary creative research laboratory at the University of Geneva. She has received several scientific and artistic awards for her work on virtual humans in Canada and in Europe. Author and coauthor of more than 200 research papers, she has directed and produced several films and mixed real-time reality shows, and created Virtual Marilyn Monroe for the film "Rendez-vous in Montreal" in 1987. She is

editor-in-chief of the Visual Computer Journal published by Springer Verlag and co-editor-in-chief of the Journal of Visualisation and Computer Animation published by John Wiley (www.miralab.unige.ch).

#### **4. Laurie McCulloch, panelist**

Ananova marked a milestone in interactive digital humans, but many others are gaining notoriety as well, including a digital William Shakespeare (who hosted BBC comedy programs), Head (a virtual talk-show host), and VIA (a range of Virtual Interactive Assistants). The trend has been to endow virtual characters with intelligence to enable them to interact more naturally with humans, thus heralding new and potentially limitless roles for digital characters.

##### **4.1 McCulloch bio**

Laurie McCulloch is development director of Digital Animation Group, developers of virtual human characters that range from newscasters and avatars to actors and models. The most famous of these is Ananova, a 28-year-old, green-haired, digital character who delivers news over the Web.

#### **5. Evan Marc Hirsch, panelist**

For better or for worse, the games industry has been working with “Virtual Humans,” or more specifically AI controlled characters, for more than 10 years. With the advent of 32-bit consoles, games began to rely less on the user’s imagination and more on the team of programmers and artists for how digital humans were portrayed. Now, the next generation consoles have arrived, and the challenges have increased exponentially. We may have more power and RAM to push more polygons and detail, but the real challenge for developers is creating non-linear characters that have depth and personality. This has had fascinating and frustrating implications for the gaming business.

##### **5.1 Hirsch bio**

Evan Marc Hirsch is Head of Visual Development for Electronic Arts Worldwide Studios and is responsible for developing interactive content for next-generation gaming consoles. Over the last year, he worked directly in Art Direction and Technical Art Direction capacity on two titles currently in production. Prior to that he was Head of Visual Development for Electronic Arts Europe, where he responsible for visuals and graphics on PS2 Launch titles. From 1996 to 1999, he served as Manager, Motion Capture, Electronic Arts Canada; where he was responsible for all motion capture requirements from pre-production to finished product for more than 25 games and supervised motion capture on 10 titles that each sold over 1 million units. He delivered a speech at the Belgium Film and Animation Festival entitled “Bringing Life Into Video Games,” which focused on the challenges and the techniques for developing characters that have depth and personality in real time games.

#### **6. Phil LoPiccolo, moderator**

##### **6.1 LoPiccolo bio**

Phil LoPiccolo is Editor-in-Chief of Computer Graphics World magazine, a position he has held for eight years. Prior to working at CGW, he served as Senior Editor of MIT’s Technology Review magazine, Managing Editor of McGraw-Hill’s Popular Computing magazine, and Executive Editor of the New England Journal of Medicine’s Massachusetts Medicine magazine. He holds an MS in Science Journalism from Boston University and a BS in Engineering from the University of Michigan.



# Extending Interface Practice: An Ecosystem Approach

Natalie Jeremijenko  
Yale Univ Faculty  
of Engineering  
nat@cat.nyu.edu

Thecla Schiphorst  
Simon Fraser Univ  
Interactive Arts  
schiphorst@techbc.ca

Michael Mateas  
CMU Computer  
Science Dept  
michaelm@cs.cmu.edu

Wolfgang Strauss  
GMD Inst Media  
Communication  
wolfgang.strauss@gmd.de

Will Wright  
Maxis Software  
wwright@maxis.com

Andruid Kerne  
Creating Media  
andruid@acm.org  
Moderator

## Abstract

Interface ecology is an emerging metadisciplinary approach, in which the creation of rich interactive experiences spans n disciplines -- such as computer graphics, mathematics, gaming, visual art, performance, and cultural theory. Interfaces extend beyond interactive artifacts, activities, and social spaces, forming intricate ecosystems. Interfaces are the catalytic border zones where systems of representation meet, mix, and recombine. Through this recombination, interface ecosystems generate fundamental innovations of form, experience, knowledge, and technology. This panel brings together a diverse range of practitioners who work from concept to experience not in terms of a particular discipline, *métier*, or medium but with a practice that interconnects multiple systems, forming a whole.

**Keywords:** interface ecology, interface, metadiscipline

## 1 Introduction

The creation of rich interactive experiences spans an n-dimensional conceptual space, in which disciplines -- including computer graphics, mathematics, gaming, visual art, performance, and cultural theory -- function as basis vectors. Interface ecology is an emerging metadisciplinary approach. In this ecological approach to interface, developers assemble diverse media (along with their epistemologies and creative, technological and semiotic processes) as expression. This gives form to human experience.

In social spaces, people's activities and behaviors are connected through interactive artifacts, processes and experiences. Interfaces extend beyond these artifacts, forming intricate ecosystems. Interfaces are the border zones where systems of representation meet, mix, and recombine. Through this recombination, signs - the semiotic units of meaning -- flow into new configurations that engage participants. The structure of interface ecosystems has a catalytic effect on these processes of recombination and engagement, setting emergent phenomena into motion. Interface ecosystems generate fundamental innovations of form, experience, knowledge, and technology.

This panel brings together a diverse range of practitioners who work within an ecological framework: *who ecologize*. Each of the panelists moves from concept to experience not in terms of a particular discipline, *métier*, or medium, but with a practice that interconnects multiple systems, forming a whole. They will discuss their processes of assemblage, and results that have emerged. They may also look to the future, and brainstorm about how the ecosystems approach will influence new interfaces.

## 2 Natalie Jeremijenko Validating Interaction with Tangible Devices

Tangible interfaces to information technologies combine digital information and physical devices. These have proven popular, mediagenic, and yet remain difficult to validate. Of the accounts

for the success of these interfaces, none provide empirical evidence or measures. I will present recent studies that examine the interactive activity with tangible interfaces, vis a vis a screen based display of the same information and purpose. The common claim, that Tangible Media is 'more intuitive' by virtue of its familiar physical form, is refuted by this data. Naive users made more errors with the tangible media devices than with the well-codified interaction strategies of screen based interaction. Nonetheless, this work does find evidence for the effectiveness of Tangible Media. Understanding how tangible media works requires the capture and measurement of the interaction through over and around, rather than with interfaces. This study presents evidence for understanding: a) the role of the physicality of tangible media; including its persistence and attention directing function; b) the role of peripheral participation, monitoring and interaction in the different contexts and applications; c) the role of shared use on interaction including: how errors are corrected and variability managed; how skill is developed; how to characterize a 'use career' of these applications; how multi-user interfaces effect the single user interaction; what aspects of the information and interaction scripts presented are legible to a wide variety of users and which are prone to more misinterpretation (or variability in interpretation); the role of open-ended interaction scripts.

I will show video data of the interactivity that is staged around the tangible devices in four studies. Building on these findings, I present several new projects that extend the potential of Tangible Media. One set of projects exploits mechanical actuation as the parameter with which to display information, and a second set of projects adapts the nonphysical strategies of tangible media to several screen-based interfaces.

Natalie Jeremijenko is a design engineer and technoartist, whose work examines how technology works, in technical and social accounts. Recently she was named one of the top one hundred young innovators by the MIT Technology Review. Her work has been presented at Tate Gallery, MASSMoCA, Rotterdam Film Festival (2000), Guggenheim Museum, Museum Moderne Kunst, Frankfurt, LUX Gallery, London, Whitney Biennial, Documenta, Ars Electronica, the Museum of Modern Art in New York, and at MIT Media Lab. She was a 1999 Rockefeller fellow. She did graduate engineering studies at Stanford University in Mechanical Engineering, and at the University of Melbourne in the History and Philosophy of Science Department. She is known to work for the Bureau of Inverse Technology.

## 2 Thecla Schiphorst Body Interfaces: Navigating Sense and State Space

Dominant western paradigms underlying the development of digital technologies have typically excluded knowledge domains of experiential body practice. Interface design can be informed by, and extended through the application of methodologies articulated within these practices. These include fields of somatics, theatre, dance, and non-western movement forms such as tai chi and martial arts. Rigorously articulated, first person or experiential

methodologies provide models for knowledge acquisition, information design, networked connectivity, remote sensing and ecological multivocality. Applications of these models suggest an ecosystem approach in which inter-activity is coupled with inter-subjectivity and inter-affectivity.

Body interfaces can share system states between multiple bodies and their multiple interactions. The answer to the question ‘What is Body?’ is an evolving, shifting construction in the arts as well as the sciences. Within interface practice, intentional grammars can be developed to intermingle meaning, presence and agency. We need to draw on our ability to dynamically map our understanding of ‘what is body’ into interface practice. We can alleviate current expressive impoverishment to extend dynamic range, by including the intimate, the intelligent, the sensory, and the taboo. I will illustrate these concepts through my work,

In *Bodymaps: artifacts of touch* (1996), the input of touch re-directs and re-positions the habits of our visual perceptual systems. This creates interplay between the liminal, sensual connections made through the direction of the body’s attention. *whisper* is a new work based on small wearable devices and handheld technologies. The *whisper* [wearable, handheld, intimate, sensory, personal, expressive, responsive] system constructs networked messages based on inferred states of the carrier bodies [which host the small wearable devices]. *whisper[s]* are wearable body architectures. Intention functions to direct and apply whispered messages, which range between direct and subliminal, suggestive and overt, seductive and definitive.

Thecla Schiphorst is a Vancouver based computer media artist, and an Associate Professor in Interactive Arts at Simon Fraser University. She is the recipient of the 1998 Canada Council biennial PetroCanada Award in New Media. Her formal education in computing science and contemporary dance has shaped her work, which integrates models of scientific representation with the experience of the physical and technical body. She is a member of the original design team that developed Life Forms, the computer compositional tool for choreography, and has worked with choreographer Merce Cunningham since 1991. She is a PHD Candidate in the CaiA-Star program at the University of Plymouth in the School of Computing, and has an interdisciplinary MA in computer compositional systems from Simon Fraser University.

### 3 Michael Mateas Expressive AI

My work is in Artificial Intelligence (AI) based art and entertainment. I simultaneously engage in AI research and art making, a research agenda and art practice I call expressive AI.

Expressive AI has two major, interrelated thrusts: (1) exploring the expressive possibilities of AI architectures - posing and answering AI research questions that wouldn’t be raised unless doing AI research in the context of art practice, and (2) pushing the boundaries of the conceivable and possible in art - creating artwork that would be impossible to conceive of or build unless making art in the context of an AI research practice.

The fusion of art and AI can be conceived of in terms of a shared interest in exploring what it means to be human, and a shared methodology of knowing-by-making. The field of Artificial Intelligence is a recent incarnation of an age-old quest or dream, the dream of building an image of the human in the machine. It is this dream, fueled by science fiction representations of AI such as Hal 9000 or Commander Data, which is the initial inspiration for

many researchers entering the field. This dream is not just about modeling rational problem solvers, but about building machines which in some sense engage us socially, have emotions and desires, and, through our interactions with them, tell us something about ourselves. AI is a way of exploring what it means to be human by building systems. An AI architecture is a machine to think with, a concrete theory and representation of some aspect of the human world. Art also explores what it means to be human by building concrete representations of some aspect of the human world. Artists often explore aspects of humanity which have been under-explored or ignored in AI research.

Combining these two ways of knowing-by-making opens a new path which takes seriously the problem of building intelligences that robustly function outside of the lab to engage human participants in intellectually and aesthetically satisfying interactions which, hopefully, teach us something about ourselves.

My presentation will explore methodological and conception issues in expressive AI, particularly the notion of a doubled system which consists of a technical machine engaging in uninterpreted computation, and a semiotic machine which organizes the rhetorical strategies used to narrate the operation of the machine. These ideas will be illustrated with example AI-based artworks, such as the interactive drama, *Facade*.

At Carnegie Mellon, Michael Mateas is adjunct faculty member in the Entertainment Technology Center, Research Fellow in the Studio for Creative Inquiry, and a PhD student in Computer Science. Michael has presented work at SIGGRAPH, New York Digital Salon, AAAI, the Carnegie Museum, the Warhol Museum, and the Walker Museum. Previously, Michael worked at Intel Labs, where he co-founded GEAR (Garage Ethnography and Applications Research), a research group employing ethnographic techniques to understand how new computing technology fits into people’s lives. Michael received his BS in Engineering Physics from the University of the Pacific and his MS in Computer Science (emphasis in HCI) from Portland State University.

### 4 Wolfgang Strauss Interfacing Mixed Reality as an Ecology of Aesthetics

To connect the notion of interface with the term ecology reminds me of a passage by Paul Virilio: "traditionally architecture and design are related to interface the exterior world to the human, the design of landscape, buildings, stages etc.; now we have to care both for design of exterior and interior spaces; those new electronic interior spaces, mainly imagery, build up the look and feel of our electronic mindscapes. They are part of the urban ecology."

Unfortunately, ecology in our built environment is usually fed by very traditional visions directed backwards. This makes life quite boring. An example arises in Berlin, a former focal point of cold war. There is a serious decision about rebuilding the traditional heart of the city. The Schloss (castle) represents the hierarchical Prussian state of last centuries. Situated on the Schloss site, and opposite the central government building of the former GDR, the Palest deer Republic (palace of republic), has been renovated at a cost of 80M Euro, due to the ecological disaster of asbestos contamination. The palace building is now finished and will be closed for demolition. Estimated costs to run the building in a provisional condition, giving space for emerging culture, is just 1.7M Euro. Politicians say: "No money, no way."

The vanishing aesthetic awareness of public spaces is beaten by digitized consciousness, somewhere on the net, in favor of a castle

in the air. What we really need are interfaces for living in mixed realities, creating ecosystems rather than constructing artifacts.

The goal of the development of [netzspanung.org](http://netzspanung.org) is an architecture for making visible the interrelations between media art, science and technology. In order to realize this, we are exploring the extension of the common notion of web platforms as means of presenting and sharing information, toward the model of an *online media laboratory*. By this, we mean a web-based platform that combines tools for contextualization of information into a collaborative knowledge space, with tools for active experimentation with networked media spaces. This takes into account the fact that the use of the web for creation and distribution of information in different disciplines (e.g. art, science, technology) is today perhaps the most significant example of mixed realities: the contents of the web represent a myriad of different perceptions of "realities", of "knowledge about" and representations of the world, expressed as networked constructs combining different media (text, image, video, 3D, mobile communications etc.) and often as a result of a collaborative process. Such a highly meditated situation of communicating and constructing knowledge requires new models for discovering contexts and relationships and for understanding how meaning is encoded in complex structures of networked media. This concern cannot be met with the "old" model of a passive user with arbitrarily intelligent" technologies. Rather, tools that enable (empower) the user to actively explore her/his own ways, and construct her/his own models for dealing with information, become essential.

Wolfgang Strauss is an architect. He has held teaching positions in Interactive Art at the HDK Berlin, at the KHM Media Art School Cologne, at the School of Fine Arts Saarbrücken and the Kunsthochschule in Kassel. He co-founded Art + Com, Berlin. Currently he is research fellow at the Fraunhofer - Institute for Media Communication. In opposition to the theory of the disappearing body, he introduces intuitive interfaces for playful interaction. His work has been presented at ZKM, Nagoya Science Museum, SIGGRAPH, ICC Tokyo, Imagina, ISEA and was awarded with the Golden Nica at Ars Electronica 1992.

## 5 Will Wright SimCity and The Sims

One of the primary roles of any designer is to fully represent the end user of a product. As a game designer I find myself not just standing in for the typical player but for all potential players regardless of their skill, interest or motivation. Many recent games are beginning to "leave the box", that is they're using web sites, player customization tools, databases of player-created scenarios and so forth to expand the scope of what users can do with the product. In a sense the players are becoming more and more an integral part of the development team. As a more diverse set of player activities becomes available the players themselves diversify to fill these niches. The different niches of player activity (tool builders, skin artists, web masters, browsers, casual players) grow to be quite interdependent and self-supporting. The game community functions as an ecosystem. As the activities around a product leave the box, so must the responsibilities of the designer. I used to think that my job as a game designer was to create a cool game, I now think my task is to facilitate the creation of a cool community with a game at the core.

Will Wright is the creator of both the SimCity and The Sims franchise. SimCity was released in 1989, and within a few months became a hit. The latest incarnation and definitive version of SimCity, SimCity 3000 Unlimited, has continued in this

tradition. Wright's game, The Sims, puts players in charge of the lives of a neighborhood of simulated people. Released in February of 2000, this title has become a cultural phenomenon. The Sims has sold over 5 million copies worldwide to become the best selling PC game of all time. The Sims has inspired four expansion packs; Livin' Large, House Party, Hot Date and Vacation. Combined sales for The Sims' franchise total 11 million units life-to-date. Next up for Wright is The Sims Online(tm). Scheduled for release in the second half of 2002, The Sims Online will enable you to take your Sims to an online world where you get to be yourself or whoever you want to be.

## 6 Andruid Kerne The Conceptual Space of Collage

*CollageMachine* [<http://mrl.nyu.edu/collagemachine>] is a creative web visualization tool that learns while you surf. Instead of waiting for you to click a hyperlink, the program proactively crawls the web, seeking content of interest. *CollageMachine* parses websites, modeling the web as collections of linked documents and their constituent media elements - images and chunks of text. These media elements continuously stream into a dynamic, user-interest-driven collage.

You can use collage design tools to create your own look and feel. By engaging in visual design, you also express dis/interest in media elements. *CollageMachine* learns about what you like from these interactions, and annotates its model to represent your interests. Decisions about what content to pursue and how to build the collage are made according to the model. The Collage Visualization Grid allocates screen real estate and history-enriches collage elements as a representation of your intentions. Unlike typical information visualization systems, perceptible structure develops bottom up. Navigational trajectories and combinatorial concepts emerge. The user experience blurs boundaries between web browsing and art-making.

The *CollageMachine* interactive artifact and the interface ecology theoretical framework are being co-developed, through back and forth loops of reference on multiple levels. *CollageMachine* promotes the emergence of new ideas through hypermedia combinations. Interface ecology, as a metadiscipline, investigates the process of combining whole systems of representation. This investigation proceeds both structurally, and in the situated contexts of particular applications, connecting theory and practice. For example, through investigation of *collage* and *emergence* -- in the context of *CollageMachine* development -- their application on the conceptual level -- in interface ecology -- became apparent. Thus, in this co-development process, theory does not inform practice simply; rather strange loops of reference, operation, and influence emerge through multiple levels of collage.

Andruid Kerne [<http://www.andruid.com>] is a research artist scientist who specializes in information visualization, agents, databases, audio, video, distributed real time systems, and public installation. He opens the range of social processes embodied by computational artifacts, for instance, substantiating play as a mode of activity and interaction. His work has been presented at SIGGRAPH, SIGCHI, the Guggenheim Museum, New York Digital Salon, ISEA, Milia, Ars Electronica, and the Boston Cyber Arts Festival. Kerne holds a B.A. in applied mathematics from Harvard, an M.A. in music composition from Wesleyan, and a Ph.D. in computer science from NYU. Andruid was recently a visiting professor at Tufts University, where he taught courses in human computer interaction, object oriented game programming, and public web installation.

# The Future of Computer Graphics: An Enabling Technology?

David S. Ebert\*  
Purdue University  
Panelist & Moderator

Bill Buxton†  
Alias | Wavefront, Inc.  
Panelist

Patricia Davies ‡  
Purdue University  
Panelist

Elliot K. Fishman §  
Johns Hopkins Hospital  
Panelist

Andrew Glassner ¶  
Coyote Wind Studios  
Panelist

## 1 Introduction

Computer graphics research and hardware has matured as a field to the point that high-quality computer graphics is becoming ubiquitous. Computer graphics shortly will be where word processing is today: everyone uses it, but there are very few people doing basic research in word processing. All of the challenges lie in the applications and use of this technology to enable advances in many fields. This panel will combine experts in computer graphics and associated technology with experts from a few applications areas to discuss the possibilities and future ways that computer graphics can advance discovery in many fields.

## 2 David Ebert

With the latest advances in graphics hardware and software technology, high-quality graphics will shortly be ubiquitous and good enough for many applications. The unsolved problems that still remain are to make smaller and smaller improvements to the existing state of the art or to improve the speed and cost of the solutions. The main challenges that I see are in applying these techniques to effectively convey information to users who are trying to solve problems and deal with the data deluge that they are facing. Combining techniques from vision, art, illustration, haptics, and, most importantly, perception are essential to solve these problems. My collaborations with scientists, physicians, and information analysts have shown me that graphics developments are most effective when combined with expertise from other areas to solve real problems.

### 2.1 David Ebert Bio

David Ebert is an associate professor at Purdue University and has been doing research in computer graphics, procedural techniques, and visualization for fourteen years. While better known in the computer graphics community for his research in simulating natural phenomena and procedural techniques (chairing several SIGGRAPH courses and co-authoring a book on this topic), his main research is in techniques for effective visualization, developing and applying techniques from volume rendering, technical illustration, and, most recently, perception to applications ranging from document analysis to astrophysics to medical diagnosis and accuracy. He has served on several ACM SIGGRAPH committees, co-chaired the IEEE Visualization papers program, and serves as Associate Editor in Chief for IEEE Transactions on Visualization and Computer Graphics.

## 3 Bill Buxton

Today we find ourselves in the interesting situation where, on the one hand, there is an explosion of resources (such as cards, game

machines, computers, net, ...) and stagnation, or worse in the industry (such as recent year to year decrease in the size of the SIGGRAPH trade show).

We now know how to do graphics and deliver them at mass/consumer level. While we have been brilliant at addressing the technical issues, we have been far less so in answering the question, "So what?" or "Where is the value?"

Where we are right now is a space where, outside of feature films and video games, the main contact with 3D for most people is in tumbling cars, running shoes, watches, etc. on the web. All of which brings to mind the immortal words of Peggy Lee's, "Is that all there is, is that all there is?" If so, as she says, "Let's go dancing," since we may have been wasting our time in a "Field of Dreams" (if you build it they will come) type scenario.

I, for one, think that there is value. But as long as the industry and research is dominated by people whose expertise is in graphics (the technology issues), rather than value (the human, social and business issue), then it will never be realized. Finally, I would argue that this realization will only come through a significant realignment of our discipline and industry.

### 3.1 Bill Buxton Bio

Panelist Bill Buxton is a designer and a researcher concerned with human aspects of technology. His work reflects a particular interest in the use of technology to support creative activities such as design, film making and music. Buxton's research specialties include technologies, techniques and theories of input to computers, technology mediated human-human collaboration, and ubiquitous computing. He is Chief Scientist of Alias|Wavefront, Inc., and its parent company SGI Inc., as well as an Associate Professor in the Department of Computer Science at the University of Toronto. While "full-time" at Alias|Wavefront, Buxton continues to supervise graduate students at the university. Buxton began his career in music, having done a Bachelor of Music degree at Queen's University. He then studied and taught at the Institute of Sonology, Utrecht, Holland, for two years. After completing an M.Sc. in Computer Science on Computer Music at the University of Toronto, he joined the faculty as a lecturer. Designing and using computer-based tools for music composition and performance is what led him into the area of human-computer interaction.

## 4 Patricia Davies

Patricia will talk about the work of a group at Purdue University in perception-based engineering, and the integration of perceptual models of human response to stimuli into the engineering design process. Applications include the design of interiors of the next generation of automobiles; and integration of visual, tactile and audio interfaces for natural and enhanced communications.

### 4.1 Patricia Davies Bio

Patricia Davies received her B.Sc. in Mathematics from the University of Bristol in 1977, and her M.Sc. and Ph.D. in Sound and Vibration from the University of Southampton in 1981 and 1985,

\*e-mail: eberrtd@purdue.edu

†e-mail: buxton@aw.sgi.com

‡e-mail: daviesp@ecn.purdue.edu

§e-mail: efishman@jhmi.edu

¶e-mail: andrew\_glassner@yahoo.com

respectively. Dr. Davies is a faculty member at the Ray W. Herick Laboratories, part of the School of Mechanical Engineering, where she conducts research in the areas of: (1) Nonlinear System Identification, (2) Vibration Measurements, (3) Sound Quality and Perception Based Engineering, (4) Acoustical Arrays for Sound Source Visualization, and (5) Event-based Signal Processing. She has applied her research to: modeling the dynamics of the polyurethane foam used in car seats, visualizing automobile noise sources during pass-by tests, predicting reciprocating pump failure, analyzing infant and mother laughter, and modeling human response to refrigerator, HVAC, chiller and large machinery noise. She has co-authored over 80 journal and conference papers. Currently Dr. Davies is also coordinating a group of psychology and engineering professors at Purdue who research and develop models of the effects of engineered products and systems on people. The aim is to integrate these models into the engineering design process so that engineered systems can be optimized for the people who use, or are affected by, them. Engineering and Psychology professors in this Perception-based Engineering group collaborate on projects such as image quality, sound quality, effects of noise on performance and decision making, speech and gesture perception, thermal comfort and air quality, touch and feel interfaces, interface design and response to vibration.

## 5 Andrew Glassner

Computer graphics has always been about building and using tools. Our tools have moved from laboratory curiosities to the local computer store's shelves, and some of our dreams from yesterday are standard components on every home system today. Science and medicine, photo editing, gaming, and film production will never be the same. What happens next?

Graphics has always been limited by the available technology, because we can only display on the hardware we've got. What happens when tabletops and wallpaper, along with the rest of the world, turn into communicating input and output devices? The whole idea of computer graphics will change from an output technique to a responsive, integrated part of our human environment.

This requires a re-thinking of what graphics is about. A satellite can gather information, but a reporter creates a story. A metronome can keep the beat, but a drummer makes it swing. Synthetic images contain data, but people give them meaning. As a field, computer graphics needs to look up from the spreadsheet and move out into the world, where we can use our tools to enhance and enrich both everyday and peak human experiences.

### 5.1 Andrew Glassner Bio

Andrew Glassner is a writer-director, and a consultant in story structure, interactive fiction, and computer graphics. He started working in 3D computer graphics in 1978, and has carried out research at the NYIT Computer Graphics Lab, Case Western Reserve University, the IBM TJ Watson Research Lab, the Delft University of Technology, Bell Communications Research, Xerox PARC, and Microsoft Research.

## 6 Elliot Fishman

When a sample of physicians was asked what the greatest impacts in medicine have been over the last few decades, they did not choose antibiotics or laparoscopic surgery or stem cell research, but instead chose medical imaging with CT and MRI.

With the newest scanners providing datasets of between 1000 and 2000 slices of 512 x 512 resolution, the challenges range from

processing data on the scanner to transferring the images to workstations to being able to load data sets into memory at sufficient speed to make the process interactive. The use of isotropic data sets is ideal for high resolution 3D imaging, but only if the interactivity for the postprocessing of the data is present. With the advances in image processing and the improvement in the quality of the data set, the generation of 3D CT maps has come a long way since the mid-80's and early 90's. Where in the past the key areas of visualization were related to soft tissue or bone, currently there is a revolution in imaging ranging from applications of the vascular system to organs such as the colon and airway. The ability to scan faster allows the acquisition of vascular angiographic maps, which are equal to and better quality than classic angiography, and yet costs one-third to one-quarter as much without any of the invasiveness. The ability to acquire these data sets is revolutionizing the practice of medicine. The ability to create vascular maps with relationships of tumor to vessels has impacted greatly on surgical planning. However, the ability to do surgical planning based on staging patients needs to go far more than where we are today. The ability for the surgeon to interact with the data set and plan the surgery interactively and then make changes on the fly in the operating room is not currently available. Although select sites have used processes to simulate surgery, they are done in a research mode or for select cases and have not really become the standard of care or even available to most patients.

The process of physician interacting with data has changed little with the exception of the willingness to use three-dimensional images. However, most centers review these images on static films and, as we have proven in the past, 3D imaging with real-time rendering is critical. Challenges to provide interactivity to referring physicians wherever they are – be it in surgery, their office, the clinic or home – needs to be met and is a challenge that goes unsolved.

The tools available to the physician have improved greatly over the years with real-time rendering. However, these tools are still, for the most part, rudimentary. Although some companies and research sites have developed new tools, they are not widely available. With the cost of medicine increasing far greater than the GNP, there is hope that computer assisted imaging and analysis of data, when done correctly, will help lower the cost of medicine and, at the same time, improve patient care. This has always been a challenge and is still an unfulfilled promise.

In this panel, I will present some of the latest images and state-of-the-art studies as seen in our practice today. I will also address what changes are needed to make these images more helpful to the referring physician and how we can better reach the ultimate goal which is to improve care for our patients. The challenges for the next decade will be defined.

### 6.1 Elliot Fishman Bio

Elliot K. Fishman, MD, FACR, is Professor of Radiology and Oncology at The Johns Hopkins University School of Medicine, and Director of Diagnostic Imaging and Body CT, Russell H. Morgan Department of Radiology and Radiological Science, The Johns Hopkins Hospital, Baltimore, Maryland. Dr. Fishman's research interests include spiral and multidetector CT scanning, CT angiography and computer-based solutions to medical imaging. He has helped develop the field of 3D medical imaging with particular interest in real-time visualization and volume visualization. Dr. Fishman has authored or co-authored more than 700 peer-reviewed scientific papers.



# Symposium on Computer Animation in Fast Forward

**Michael Gleicher**, University of Wisconsin - Panel Moderator

**Michael Cohen**, Microsoft Research - Symposium Papers Co-Chair

**Nancy Pollard**, Brown University - Symposium Papers Co-Chair

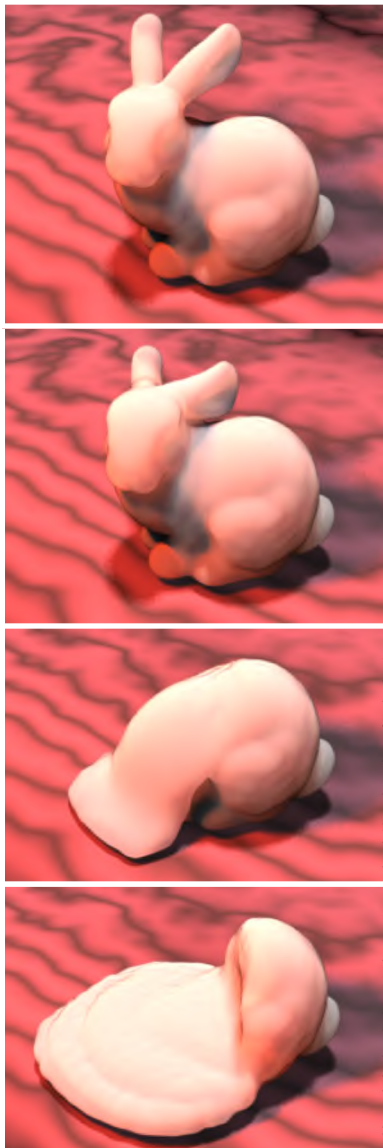
**Jessica Hodgins** - Symposium Co-Chair

**Michiel van de Panne** - Symposium Co-Chair

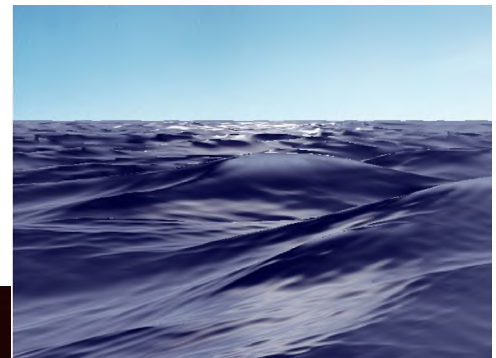
"Three-minute madness." Papers from the new ACM SIGGRAPH Symposium on Computer Animation are summarized in three minutes or less, followed by a discussion of new directions in computer animation research.

The ACM SIGGRAPH Symposium on Computer Animation (SCA) was created to provide a new venue for discussing ideas relating to all aspects of computer animation.

The first Symposium (SCA2002) was held July 21-22, colocated with the SIGGRAPH 2002 conference. Twenty two papers were presented. In this panel, each author is provided an opportunity to summarize their work to the SIGGRAPH community.



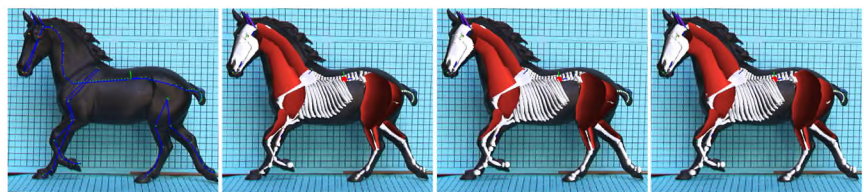
Melting and Flowing. Carlson, Mucha, Van Horn, and Turk (Georgia Institute of Technology).



Interactive Animation of Ocean Waves - Hinsinger, Neyret, and Cati (IMAGIS-GRAVIR/IMAG-INRIA)



Motion Capture-Driven Simulations that Hit and React. Zordan (Georgia Tech) and Hodgins (CMU)



Automatic Model-Based Reconstruction for Creature Animation. Simmons, Wilhelms, and Van Gelder (University of California, Santa Cruz)



# The Demo Scene

Vincent Scheib  
UNC Chapel Hill  
Demoscene Outreach Group  
scheib@cs.unc.edu

*Organizer/Moderator*

Theo Engell-Nielsen  
hybris/NEMESIS  
theo@rhk.dk

Eric Haines  
Autodesk, Inc.  
erich@acm.org

Saku Lehtinen  
Remedy Entertainment  
saku@remedy.fi

Phil Taylor  
Microsoft DirectX  
ptaylor@microsoft.com

## Abstract

For 20 years, an underground movement has produced short real-time animations running on home computers. This group, the "demoscene," primarily consists of students who pursue their technical and artistic interests beyond the classroom, to create inspiring works of real-time art. These productions encompass a broad range of computer graphics techniques such as procedural geometry, real-time ray-tracing, and real-time shading.

Game developers have been utilizing this talent pool yet it has little visibility in the SIGGRAPH community. This panel explores the demoscene, technical tricks used in demos, and how scene educational and creative aspects can contribute to the SIGGRAPH community.



Screen shot from *Moral Hard Candy*, by Blasphemy

## Demo Scene?

The Demo Scene consists of hobbyists who create Demos. A Demo is an executable program that produces, in real time, engaging computer graphics synchronized with music. Demos run on home computers – they are a celebration of the abilities of consumer equipment and the developers' skill. Programming, art, and music composition skills are stressed. Demos are similar in some ways (but are not equivalent) to music videos or short films.

Over 100 'parties' take place per year at which Demo competitions are held. In Europe, several annual parties run for 4 days and host as many as 6,000 guests. Besides the challenge of real time graphics algorithms, Demos are also often restricted in file sizes such as 4KB or 64KB. This requires the use of

procedural textures, geometry, and animation. Graphics programming alone, however, is not sufficient. Demos also contain songs using elaborately synthesized instruments, 2D & 3D artwork, and key framed & motion captured animation.

## Theo Engell-Nielsen – Active Scener

Theo has a Master of Science in Computer Science & Information Psychology, and is a Ph.D. student at the IT University of Copenhagen, Denmark. He is the founder of the demo group called "hybris/NEMESIS" (<http://www.hybrismnemesi.com>).

Theo has been a demo coder since the mid 80's on the Commodore 64 (assembly), Amiga 500 (assembly), PC platform (MS Windows, C++), has made scene related music on the Amiga 500 and is now a member of the WILD!Demo demo group hybris/NEMESIS.

He acts as organizer, originator of ideas, technical director, art director, programmer, modeler, animator, sound fx, ... the works. He is an experienced system programmer (2 years), multimedia programmer (4 years), games programmer (1 year), 3D modeler, and 3D animator. Finally, he plays in the Commodore 64 revival band PRESS PLAY ON TAPE (<http://www.PRESSPLAYONTAPE.com>).

## Eric Haines – Real-time Graphics

Eric Haines has been involved in SIGGRAPH since 1984, and has been involved as a researcher, a contributor to the art and film shows, and a presenter at a number of courses. He is a coauthor of "An Introduction to Ray Tracing," and has helped disseminate information on this topic to the research and hobbyist community for many years through the electronic magazine "The Ray Tracing News" and the "Ray Tracing Roundtable" at SIGGRAPH.

More recently he coauthored "Real-Time Rendering", released this summer in its second edition, and runs a popular informational site about interactive graphics techniques. He is also an editor of the "Journal of Graphics Tools" and the maintainer of the "Graphics Gems" code repository and the ACM TOG site. He received a Masters from the Program of Computer Graphics at Cornell. He currently works for Autodesk, Inc..

## Saku Lehtinen – Game Industry

Saku works as a graphic and game designer at Remedy Entertainment. Remedy's most known work is a game called Max Payne. He has also been involved with the development of technology used at Remedy, the Max-FX, and especially the level editor MaxED. Saku teaches multiple courses in the Architecture department at the Helsinki University of Technology including topics on CAD, DTP, visualization, web, and design. He was part

of the group Aggression in the early 1990's, which was active in the Atari ST scene.

### **Phil Taylor – DirectX**

Philip Taylor is the project manager for Microsoft's DirectX SDK, Managed DirectX, the WindowsXP 3D screensavers, and a few more bits and bobs. Previously at Microsoft he was senior engineer in the DirectX evangelism group for DirectX 3.0 to DirectX 8.0, and helped many game ISVs with DirectX. He has worked with DirectX since the first public beta of the GameSDK (DirectX 1.0), and, once upon a time, actually shipped DirectX 2.0 games. In his spare time, he can be found lurking on many 3-D graphics programming mailing lists and Usenet newsgroups. His interest in the scene is natural considering his long interest in cool 3D.

### **Vincent Scheib – Organizer/Moderator**

Vincent Scheib is pursuing a PhD at UNC Chapel Hill researching real-time interactive techniques, with academic publications in SIGGRAPH and Eurographics. Vincent has worked previously in the game and engineering industries.

He founded the Demoscene Outreach Group in 2000 specifically to bring the demo scene to the SIGGRAPH conference. Since then, he has made presentations at the Game Developer's conference, the Assembly Demo Party in Helsinki, on TechTV's show the Screen Savers, and in an article on Gamasutra.



**Screen shot from VIP2, by PoPsY TeAm**

### **More Information**

Please find more information on the Demoscene Outreach Group's web page: <http://www.scene.org/dog>

Additionally, several PC demos can be found on this year's supplementary CD-ROM.

## Animation's Turning Tide

Scott Elson  
Walt Disney Feature Animation  
Moderator

Eamonn Butler  
Walt Disney Feature Animation  
Speaker

Scott Clark  
Pixar Animation Studios  
Speaker

Carlos Saldanha  
Blue Sky Animation Studios  
Speaker

This panel seeks to demystify, debunk and goad the dialogue of 3D character animation. This panel is motivated by the sea change that is currently effecting our industry. In the last year and a half what was once a battle to get 3D into Hollywood has completely crossed over – Hollywood can't get enough. 11 years ago the quote from top Disney management was “there will never be a digital character in a Disney feature film.” Now traditional animators and artists are giving up their fear for pragmatism and embracing 3D in droves. What does that mean for those already working in 3D? What does it portend? What can we learn from the traditional animator?

At the time this panel was proposed, in the fall of 2001, “Shrek” was a hit and “Final Fantasy” a flop; was “Shrek” a fluke or a harbinger of things to come? Since then “Monsters, Inc.” and “Jimmy Neutron” have done huge box office and in the week this goes to press “Shrek” has won the first Academy Award™ for an animated feature film. By the time of ACM SIGGRAPH 2002, “Lilo & Stitch” and “Spirit” will both be out at the box office, both are basically 2D films. What will those films mean for the future of both 2D and 3D films? Is the medium the message? Hold onto your seats, it's going to be an entertaining afternoon.

# Unsolved problems in Mobile Computer Graphics and Interaction

Mark Ollila\*  
Norrköping Visualization and Interaction Studio  
Linköping University, Norrköping, Sweden  
Moderator

Steven Feiner§  
Department of Computer Science  
Columbia University, New York, USA  
Panelist

Staffan Björk†  
PLAY, Interactive Institute  
Göteborg, Sweden  
Panelist

Kevin Bradshaw‡  
Machines That Go Ping!  
Edinburgh, Scotland  
Panelist

Kari Pulli¶  
Nokia Mobile Phones  
Oulu, Finland  
Panelist

## 1 Introduction

With the number of mobile devices exceeding PCs, research is required in many areas with respect to graphics and interaction. There are problems with interaction, streaming, graphics algorithms, bandwidth with current and future devices. This panel examines the state of the art from both an industrial and research point of view, and provides directions for future work in this area.

### 1.1 Questions and Issues

The panel takes experts from industry and research and examines the issues that face the SIGGRAPH community worldwide in the area of Mobile devices. The problems requiring further research and exposure and are for debate include, desktop vs mobile; adaption of multimedia; power, battery vs electric grid; network connectivity, types of band and size of bandwidth; display, millions of pixels vs thousands of pixels; performance, high power chips vs slower chips; context of use, user is mobile; standards and formats; rendering at terminal vs rendering at server; augmented reality; location independence vs location dependence; interaction techniques; cross cultural interaction; cross network connectivity, WAN v LAN v PAN; computer graphics software design for mobile devices, application development.

The panel has representatives from industry that can provide real world case examples of what problems they have experienced, from manufacture, to human factors.

## 2 Mark Ollila

Mobile devices are becoming more and more popular world wide, as people carry PDA's, wireless phones, or portable game devices. These devices are becoming more and more ubiquitous and the implementation of applications and various ways to interact are becoming more complex. From the perspective of unsolved problems, these include: dynamic multimedia adaption algorithms; hybrid representations of multimedia data which takes into consideration terminal device and bandwidth; adaptive software components that create applications that reside on the server; context aware agents, that resides on both the client and on the network.

There are many problems that need to be addressed in terms of rendering on a battery limited device with low screen resolution. How do we deal with these problems? How do we use the screen

space more effectively? How do we handle interaction with a device when the context that the user will be in is not an office like environment.

### 2.1 Moderator Ollila bio

Mark has conducted research in many areas of Computer Science - specifically Mobile Technologies, Computer Graphics and Interaction with a touch of Artificial Intelligence. He has worked in administration, programming, new media. He loves innovation and thrives on taking research to industry. An active person, Mark has setup research labs, studios and startups around the world. After leaving Australia to discover his Scandinavian heritage seven years ago, Mark was given the position of Head of Creative Programming at the University of Gävle. Mark worked at the Interactive Institute in Stockholm where he helped developed the strategy proposals for the newly started Mobility Studio. Currently, Mark has a position at Linköping University in building up a Media Technology Institute. He also has a position at the MediaTeam at the University of Oulu, focussing on MobileServices. His current research interests are P2P mobile networks, client side personalization and privacy, emotional artificial intelligence, image based rendering solutions, and theory of gaming. He is a representative for Sweden in the World Internet Project, specifically on Mobile related issues. Mark has a PhD in Computer Vision. He has been on the international committee for SIGGRAPH for 1998,1999,2000,2002.

## 3 Staffan Björk

The high availability of cheap portable computers combined with wireless communication (via WLAN or mobile phone networks) has increased the interest in various forms of mobile and networked applications. However, most applications try to enable the same functionality that exists on stationary computers without taking into consideration the changed use context. How should one design an interface that is to be used in an environment such as a subway tram where the surrounding light varies? Where the user is constantly being shook? Where the application cannot require the user's attention as he or she is talking with other people? How to design applications that are part of people's lives rather than their work?

Designing applications in this new setting requires a different view on what computer interaction is. Functionality will have to coexist with aesthetic appeal and applications will be defined by how they support the user's activities in the real world rather than what file formats they support in the computer.

### 3.1 Panelist Björk Bio

Staffan Björk is head of the PLAY research group at the Interactive Institute and has conducted research on Information Visualization

\*e-mail: marol@itn.liu.se

†e-mail: staffan.bjork@interactiveinstitute.se

‡e-mail: kevin@mtgp.net

§e-mail: feiner@cs.columbia.edu

¶e-mail: kari.pulli@nokia.com

since 1998 together with Ericsson Usability Lab, Telia Research and Nokia Research Centre. He received a PhD in the autumn of 2000 for his work on the WEST, PowerView, and PowerCom interfaces for small displays. He has concurrently done research within the field of Ubiquitous Computing, most noteworthy Pirates! which was exhibited as part of the Emergent Technologies section of SIGGRAPH 2001.

## 4 Kevin Bradshaw

Mobile/wireless companies today have a desire to create and deliver real-world, usable and effective interactive digital media applications to wireless devices - devices limited in screen size, user interface, bandwidth, storage capability and power. Despite these limitations, these devices are the most interesting of all areas for research into interactive media design and delivery and they will number well over one billion active users within 4-5 years.

Major problems have been encountered to date even throughout the short lifetime of availability of advanced wireless media terminals. The wireless industry did a good job of standardising and bringing to market interactive text (SMS) terminals, but made a very poor initial job of launching interactive browser terminals. The launch of such devices and services was plagued by incompatibilities amongst vendors, greatly varying implementations of UI standards and thus almost insurmountable barriers to entry for new media developers. This has led to a very troubled birth for interactive wireless media services.

### 4.1 Panelist Bradshaw Bio

Kevin Bradshaw founded Digital Bridges in 1998, with the aim of developing a global distribution technology and a set of high-quality entertainment channels for Wireless Internet Users. Prior to Digital Bridges, Kevin spent four years working as a consultant for Consult Hyperion, one of the world's leading electronic and mobile commerce innovators, and the Smith Group, an international Science and Technology consulting organization. Kevin is currently the Chair of the Advisory Board for Digital Bridges and also a Director of the Mobile Entertainment Forum, the first global trade association for Wireless Entertainment. Kevin is the Founder and CEO of MTGP, a strategic advisory and investment group for ventures in Interactive Media Technology. Kevin received his PhD. from Oxford, UK.

## 5 Steven Feiner

With the latest laptops now packing more 3D graphics power than high-end graphics workstations of a decade ago, it's clear that far more powerful and far smaller mobile devices will follow. One form will be hand-held: phones, PDAs, and gaming platforms. However, another will be head-worn: see-through eyewear, ultimately head-tracked, overlaying graphics on the real world to create augmented reality. Many technical problems remain to be solved before accurately registered, wide field-of-view, sufficiently bright, comfortable and attractive, augmented reality displays will be commonplace. Independent of these technological issues is the question of what we will want to see overlaid on our world. I will suggest some of the possibilities, ranging from nothing (Can we really improve on a perfect sunset?), to satisfyingly minimalist (a single arrow pointing to the next button to push in a maintenance procedure), to breathtakingly photorealistic (experiencing today the Forum Romanum of two millennia ago). Along the way, I will discuss why many of the CPU cycles in these future devices will be spent designing what we see, not just rendering it.

## 5.1 Panelist Feiner Bio

Steven Feiner is a Professor of Computer Science at Columbia University, where he directs the Computer Graphics and User Interfaces Laboratory. He received a Ph.D. in Computer Science from Brown University in 1987. His research interests include virtual environments and augmented reality, knowledge-based design of graphics and multimedia, information visualization, wearable computing, and hypermedia. Prof. Feiner is coauthor of *Computer Graphics: Principles and Practice* (Addison-Wesley, 1990) and of *Introduction to Computer Graphics* (Addison-Wesley, 1993). He is an associate editor of *ACM Transactions on Graphics*, has served on the executive boards of the IEEE Computer Society Technical Committee on Visualization and Graphics and the IEEE Computer Society Task Force on Human-Centered Information Systems, and is a member of the steering committees for the IEEE Symposium on Information Visualization, the IEEE and ACM International Symposium on Augmented Reality, and the IEEE International Symposium on Wearable Computers. Over the past year, he has been general chair of IEEE Information Visualization 2001, symposium co-chair of the IEEE and ACM International Symposium on Augmented Reality 2001, and program co-chair of the International Symposium on Mixed Reality 2001. In 1991 he received an Office of Naval Research Young Investigator Award.

## 6 Kari Pulli

Hardware architectures, power, performance, software architectures, and interaction issues are big questions facing manufacturers looking to use computer graphics. The desktop paradigm that current computer graphics seems to focus on is not going to be the same one that is used for mobile technologies. Mobility enables something more than a desktop system does: the device can be always with you where ever you go, and the services and the imagery can be tailored to your current location and context.

In future research, many factors need to be considered, and new architectures have to be studied and developed, or old ones adapted, to deliver the usability and services that end-users demand. This covers all aspects of computer graphics, such as screen technology, software architectures and APIs, as well as interaction techniques.

### 6.1 Panelist Pulli Bio

Kari Pulli is a Principal Scientist at Nokia Mobile Phones where he heads research activities that ensure mobile devices becoming visually interesting communication systems, from the input (cameras) to the output (displays and graphics). After receiving his PhD in 1997 from University of Washington in Seattle, Kari worked at Stanford University, heading the technical aspects the Digital Michelangelo project. His current academic affiliation is at University of Oulu, where he teaches computer graphics. Kari's previous industrial experience stems from stints at Microsoft, SGI, and Alias|Wavefront.

# Graphics in the Large: Is Bigger Better?

Loren Carpenter\*  
Pixar Animation  
Studios  
Panelist

Brian Fisher†  
University of  
British Columbia  
Panelist

Richard A. May‡  
University of  
Washington  
Panelist

Norbert Streitz§  
Fraunhofer IPSI  
Panelist

David J. Kasik¶  
Boeing  
Moderator

## 1 Introduction

The world of display devices is expanding rapidly, both literally and figuratively. New commercial and research devices come in larger sizes (measured in meters, not inches) and different physical forms (e.g. rectangular surfaces, cylindrical segments, truncated spheres). Such expansion means that graphics and interactive techniques are becoming far more amenable to group activities and can display more and more data at once.

This panel considers graphics in the large. We will examine large-sized graphics that can be displayed on multiple devices, devices that can be used by groups of co-located and distributed users, and the impact of displaying more and more data. The discussion will center on “Is Bigger Necessarily Better?”

### 1.1 Questions and Points for Debate

The motivation of this panel is recent work evaluating the impact of large display devices, rendering techniques, and interaction techniques (both single user and groups) for such devices.

The move toward large display devices parallels the information explosion fueled by pervasive, low cost computing. People continue to search for better ways in which they can better understand and use the huge amounts of information. The visual channel is one of the most obvious because the human vision system can accept and process pictures with amazing efficiency.

Larger and larger display devices are a natural step: people can understand more complexity and relationships when they see more. Potential applications abound because large screen devices can be used to display information in 2D or 3D. Complex 2D displays are common in command and control systems that manage power grids, computing networks, air traffic, and battlefield scenarios. The gaming and entertainment industries generate large numbers of 2D and 3D visual environments that millions view daily. Use of 3D has become commonplace. Companies that build physical products (e.g. automobiles, appliances, aerospace vehicles, buildings) rely on solid and surface geometry as the product data definition master. Display of the information in 3D has become essential in understanding geometry form, fit, and function. Information visualization techniques are invaluable for people who try to understand and analyze vast quantities of seemingly unrelated data.

The computer graphics community has made significant progress not only in large screen devices but also in graphical rendering techniques, particularly in enhancing the illusion of three dimensions. The techniques have been particularly effective for images projected onto large screens. Work still is being done to make three-dimensional images appear ‘better’.

A significant body of questions arises with the advent of larger and larger display devices. For example:

- In what situations are such devices beneficial, if any?
- What techniques can be used to prevent information overload?
- Are there ways for individuals to interact other than mouse and keyboard? How do we transition from a WIMP interface?
- Is there an effective way of allowing a group of users to interact with a single large display during a design or a review session?
- How does an individual work with multiple large devices that have different characteristics?

## 2 Loren Carpenter

Cinematrix systems have been installed in theme parks and millions of people have played games together. Cinematrix has also been extensively employed in the corporate theater market, energizing meetings, providing the actual *experience* of teamwork, and real-time statistical feedback to management. We have been studying what goes on ever since 1991 when the audience erupted. Rachel Carpenter's M.S. thesis describes the anthropological view of the audience experience. Over the past seven years we have come to understand some of the principles in effect during a computer-mediated collective group activity.

This panel's focus on large screen displays brings in Cinematrix because our events require a large screen to focus the attention of a large group. Attention focused on a common screen is needed for the synchronization that raises the energy in the room. It's precisely the difference between 500 people going to a movie theater or 500 people watching a movie on their TVs at home. Excitement and all the other emotions are contagious and non-linear with density, thus amplifying the experience.

One of the things we have discovered is that, in general, people *want to help*. There are deep historical reasons for this that will be described in detail. Also, there are cognitive elements that are essential to minimize the audience's learning curve. Finally, there are momentum and showmanship effects that need to be understood, or the audience will lose interest and drift away. Near and far future possibilities will be explored.

### 2.1 Panelist Carpenter Bio

Loren Carpenter is Senior Scientist at Pixar Animation Studios; and, with his wife Rachel, a founder of Cinematrix Interactive Entertainment Systems. He received the third SIGGRAPH Achievement Award for his image synthesis and display algorithms. In 1996, he was named an ACM Fellow. In March 2001, he received an Oscar for his pioneering efforts that led to the digital filmmaking revolution.

## 3 Brian Fisher

The choice of display environment must take into account the interaction between the proposed environment, the graphical characteristics of the images presented, and the performance

\* e-mail: loren@cinematrix.com

† e-mail: brian.fisher@acm.org

‡ e-mail: rmay@hitl.washington.edu

§ e-mail: streitz@ipsi.fraunhofer.de

¶ e-mail: david.j.kasik@boeing.com



characteristics of users' visual and motor control systems. In addition to drop-offs in acuity and color perception in peripheral vision, less well-known findings from research in spatial cognition (users' consciously accessible representation of visio-spatial information) and functional and apparent space constancy (their ability to compensate for changes in point of view for action and scene understanding respectively) are also important. Much of this is explained in the "two visual systems hypothesis" of Trevarthen that postulates neuroanatomically distinct parallel streams of processing in vision that operates differently.

From this perspective, large format displays have advantages over smaller displays only in certain situations. For example, the two visual systems hypothesis predicts that visual context illusions will have a larger impact on users in decontextualized (or large screen) immersive environments, particularly if users' interaction is determined by where they believe the display items to be (versus, for example, selecting them by pointing). Images that generate these illusions may be better viewed on a smaller screen. Of course, smaller displays are not necessarily better. Mental mapping analysis suggests situations where the ambient space constancy mechanism could better help users build a mental model of a large data space when sufficient peripheral information is present, i.e. in a large screen or immersive display environment.

### 3.1 Panelist Fisher Bio

Brian Fisher, Ph.D. (Experimental Psychology, University of California) is the Associate Director of the Media and Graphics Interdisciplinary Centre and Adjunct Professor in Computer Science, Commerce, and Psychology at the University of British Columbia. He serves on the UBC Brain Research Centre and the Institute for Computing, Intelligent and Cognitive Systems.

## 4 Richard May

Development of computer systems to support co-located collaboration has been gaining support in the research community for many years. However adapting the commonly used interface tools (mouse, keyboard, and desktop display) to support co-located collaboration is not possible. These interface devices were designed to support single user interactions, not the interactions of groups. These tools also focus the user's attention on the interface and away from other users or physical devices that are necessary for solving complex real-world problems. Additionally, these tools only support a limited (point, click, type) set of interactions that must be rigidly and uniformly applied.

The problem with designing group interfaces is an issue of design and configuration, not an issue of technology. Many technologies available today can be used to create innovative display and input devices. However, even a large-format computer screen with the resolution of printed material does not solve the problems of group interfaces. It is my position that group interface design would be greatly facilitated by applying interaction metaphors from the physical world for electronic data manipulation and by supporting situation specific interactions.

### 4.1 Panelist May Bio

Richard May is a final-year Ph.D. candidate while also maintaining his status as a senior researcher at Battelle. May's Ph.D. research is the definition, development, evaluation, and optimization of a group, direct-interaction environment. The environment uses video cameras to recognize both people and objects in the interaction space.

## 5 Norbert Streitz

My position on the issue of large screen displays results from the concepts and experiences with the "Roomware®" components we have developed over the last six years ([www.roomware.de](http://www.roomware.de)). By Roomware®, we mean computer-augmented room elements like doors, walls, furniture integrated with information technology. They are part of our approach that the "world around us" is the interface to information and for the cooperation of people. It requires an integrated design of real and virtual worlds augmenting reality. The computer disappears as a device and is almost "invisible," but the functionality is ubiquitously available. Thus, the roomware approach moves well beyond standard desktop environments.

One of many examples is the DynaWall®, an "interactive electronic wall" that is 4.5 m wide and 1.1 m high. Based on touch-sensitive screens, you write or draw directly on it with fingers or a regular pen and interact via gestures in a modelless interface using incremental gesture recognition. Due to the large dimensions, we developed new forms of human-computer interaction, e.g., "take and put" or "throw" and "shuffle" in order to move information objects from one side to the other.

Application scenarios include various types of supporting brainstorming and teamwork. For example i-LAND provides an interactive landscape for creativity and innovation. We have expanded these concepts to various other locations in the building in "Ambient Agoras: Dynamic Information Clouds in a Hybrid World" ([www.ambient-agoras.org](http://www.ambient-agoras.org)), part of the EU Disappearing Computer initiative. ([www.disappearing-computer.net](http://www.disappearing-computer.net)).

### 5.1 Panelist Streitz Bio

Norbert Streitz received an M.S. in Physics, a Ph.D. in Theoretical Physics (both University of Kiel) and a Ph.D. in Cognitive Psychology (Technical University Aachen). He is now head of research "AMBIENTE - Workspaces of the Future" of the Fraunhofer Institute IPSI in Darmstadt, Germany.

## 6 David J. Kasik

The problems of aerospace vehicle design stress the limits of a person's ability to understand complexity. The general rule of thumb has been that users can always understand any picture, especially one derived from a rendered 3D model, because the computer will run out of processing capacity before the person does. Compute capacity for rendering 3D models has increased substantially and makes graphic information overload possible.

The graphics industry has been producing some fascinating large display devices. With the increased compute power, we can now display larger parts of an airplane on a big screen. However, there has been little consideration whether users actually benefit or comprehend the increased amount of graphical information.

The graphics community can assess the effectiveness of products and interfaces in specific situations more systematically and earlier. We pay insufficient attention to interactive techniques and simplified graphical design. Instead, we often try to use a sledgehammer when a paintbrush would be more appropriate. Early assessments will drive better products for the industry and improve the ultimate usefulness of graphics for complex tasks.

### 6.1 Moderator Kasik Bio

Dave Kasik is a Boeing Technical Fellow and the Geometry and Visualization architect for Boeing Commercial Airplanes in Seattle WA. He has worked extensively in user interface software. He received an M.S. in Computer Science from the University of Colorado and a B.A. in Quantitative Studies from the Johns Hopkins University.

# How Does Motion Capture Affect Animation?

Margaret S. Geroch\*  
Wheeling Jesuit University  
Panelist

Evan Hirsch†  
Electronic Arts Worldwide Studios  
Panelist

Joan Staveley‡  
VP Faust Logic  
Panelist

Tom Tolles§  
hOuse Of mOves  
Panelist

Barb Helfer¶  
ACCAD  
Moderator

Suba Varadarajan||  
ACCAD  
Organizer

## 1 Introduction

The application of motion capture in the movie industry has continued to increase in the last couple of years, ranging from background action to major characters. Using motion capture in a production pipeline requires both motion capture and animation experience. An animator needs an understanding of biomechanics and how the body moves both for planning marker placement and to be able to apply motion capture to a character in such a way that it fits the character and the story development. For motion capture to be usable, it requires planning well in advance, knowing what needs to be captured and how it's going to be applied.

Although no one really believes that motion capture will replace the need for animators, we are interested in how motion capture has affected the animator's job. Has it made it easier or more difficult? Is it just another tool for the animator to use? This panel will be discussing the advantages and disadvantages of currently using motion capture in the production pipeline and what advances need to be made to improve motion capture usability.

## 2 Margaret S. Geroch

I believe that motion capture technology can have a much greater impact on animation and film than simply as an automated method for obtaining realistic movement in computer generated or computer enhanced characters. Through motion capture we are able to see and record in detail the visual facets of a complex human movement. A character that shares many attributes with the person who provided the motion to be captured can then be animated with the subtle details that give individuality along with realism. Much of the "talent" providing the motion, however, is professional. Actors, dancers, sports figures – those for whom movement is a professional matter are providing the action for all kinds of characters. The result is that the characters also look like professional movers, rather than like the normal people we might encounter in our daily lives. However, through this technology, in particular, by capturing the motion of many different kinds of people, both "talent" and people off the street, and thoroughly analyzing and comparing this motion, we can progress further towards understanding the range of motion attributes that people are capable of. And we can begin to identify those details which make human motion so complex and so difficult to simulate realistically.

Margaret Geroch is an Associate Professor of Computer Science at Wheeling Jesuit University and a Ph.D. Candidate in Computer Graphics, The Ohio State University and the Advanced Computing Center for the Arts and Design. Her dissertation topic involves

\*e-mail: geroch@wju.edu

†e-mail: ev@europe.ea.com

‡e-mail: jstaveley@visi.com

§e-mail: tomt@moves.com

¶e-mail: barb@cgrg.ohio-state.edu

||e-mail: varadara@cgrg.ohio-state.edu

using motion capture data to understand the range and variation of human motion for animating human figures realistically.

## 3 Evan Hirsch

Having ran a motion capture studio and worked in various capacities on key frame animation projects, I can say with great certainty that given a choice, both approaches have their time and place. For complex sports, stunts and the like, motion capture is hard to beat for rich, detailed secondary animation. We have seen numerous games and even a feature length film that were motion captured yet in many cases, when the images make their way to the viewers' eyes, something still seems slightly disturbing. On the other hand, what do you do when you need to create unique personalities for 30 complete characters in less than 9 months? Key frame animation is hardly the solution assuming your employer or client chooses to stay in business. Procedural animation could be the solution, but do we really want programmers animating our characters? Why is it so painful to edit or alter motion capture moves? Why is layering motion still so hard? The reality is that despite the hype, the tools have let us down entirely and we still have a long way to go. We've been at this for almost 10 years and yet, its still much harder then it should be. I plan to propose some principles for which approach makes sense when, and to lay out a challenge to tools developers for the perfect animation system.

Evan Hirsch was most recently based in London as Head of Visual Development for Electronic Arts Worldwide Studios where his focus was primarily on the Art Direction and CG processes EA uses to develop their next generation console games. Prior to his move the UK, was the Manager of EA's Vancouver Motion Capture Studio. While at EA Canada, he supervised Motion Capture shoots with Tiger Woods, World Championship Wrestling and a complete F1 Pit Crew and oversaw the design of the world's largest purpose built motion capture stage. Before joining EA, he was a Partner for at Acme Animation Group, a small, 3D shop that catered to Design and Advertising clients. He is a member of BAFTA, SIGGRAPH and the IDSA in addition to being a Contributing Editor to Computer Graphics World since 1990.

## 4 Joan Staveley

Motion capture and key-frame animation are two separate animating techniques. Both require skilled animators with creative vision. Each technique lends itself to a different kind and quality of expression. Motion capture is in its infancy compared to key-framing, and so far it has proven to be a good match when the creative direction requires realistic human motion. Key-frame animation is superior to motion capture as a technique when the creative direction requires the imitation of traditional animation and animation narratives. Neither technique alone determines "a great work" of animation. The factors that lead to great works are a complex set,

involving creative ingenuity, technical mastery and ingenuity, funding, and hitting a chord within society.

I arrived at the above conclusions regarding "motion-capture vs. key-frame" animation as the result of having worked at Windlight Studios in Minneapolis, Minnesota back in the 1990's. Motion capture was a new technique at the time. At Windlight we experimented heavily with motion-capture animation, attempting to use it for both realistic human motion, and for character animation. One of the beliefs that Windlight was founded on turned out to be incorrect: that motion capture could be used as a fast and inexpensive substitute for traditional character animation. Making motion capture mimic traditional animation requires a huge amount of additional animation and tweaking of the underlying motion capture curves. And, of course, one must start with a great character actor for the motion talent. Once Windlight animators developed techniques for mimicking character animation with motion capture, we produced many visually appealing character animations, but it was not cheaper or faster than key-framing the same motion.

By contrast, animating realistic, human movement using motion capture worked quite well. It was fast, looked better than key-framing in most cases, and was much more cost efficient. Most of the additional animation is put on top of the motion curves and does not require major overhauls in the timing of the underlying curves. Cost efficiency is relative, however. If a company is doing a lot of realistic human motion it makes sense to invest in setting up a motion capture studio. It is not cost efficient to invest in motion capture if there's isn't anyone on the staff who knows when and how to use it effectively, or the company does not intend on using motion capture regularly. A motion capture director must know how to pull out a great performance from an actor, and must know what kinds of motion will translate well into animation.

What remains the biggest challenge for creating realistic, human-like figures regardless of whether one uses motion capture or key-frame animation is the lack of sophistication in solving the very complicated movement of muscles, tendons, hair, skin and clothing for the body. These huge feats have to catch up with the sophistication of the skeletal motion that animators create. Realistic, beautiful people are very difficult to mimic in animation, as are realistically fat or sinewy people. The grotesque, strange and surrealistic works best because the technology for moving muscles, skin and clothing inherently communicates something "off" or grotesque. Our eyes are great at picking up on the inconsistencies in human motion. Creating more realistic-looking humans in the near future will not be dependent on either motion capture or key-frame animation. Instead, creating a closer human likeness in animation will depend on developing more realistic simulations of muscle, tendons, hair, skin, and clothing.

Motion capture has a creative future outside of animation in the performing arts. I can't wait to see dances and theater productions that make sophisticated use of this new creative tool. It has already established itself as a creative technique in animation, films and electronic games, and will continue to evolve in these media also. Motion capture, like traditional animation is here to stay!

Joan Staveley is currently Vice President of Faust Logic, Inc. in Minneapolis, Minnesota. Joan has extensive experience directing both key frame and motion capture animation for commercial and artistic purposes. Over the past ten years Joan has served on numerous Siggraph panels and juries, and has had several animations screened at the Siggraph Electronic Theater.

## 5 Tom Tolles

The good news is that motion capture is finally starting to come out of the dark ages. By a combination of improved capture hardware, capture and processing software and industry experience, certain types of motion capture work are now fairly rote. However, there

are many areas for improvement, it is a little daunting. Some typical examples include:

- 1) First unit friendliness - currently, the idea of having mocap technology on a live shoot is terrifying. Yes, with 18 or 24 cameras, we can capture quite well but losing calibration in an environment like that is far too easy to do. Solutions may include some sort of self-healing or auto-tracking calibration.
- 2) Entire Performance Capture - full body capture is pretty good these days but combined with facial and/or hand capture, all you've got is a mess. This will be real tough one to solve.
- 3) While improvements have been made in the processing of optical motion capture data, it is still far too hard to deliver initially clean capture data based on the original performer. There will be some fairly significant advances here this year, I believe.
- 4) Post-processing software - just like traditional animators study human motions to help create content, I would like to see a 3D animator be able to gracefully use motion capture data and make changes to it at a high level (hmmm, walk a bit more aggressively or fall down more like a cartoon character).
- 5) We need a combination of really easy to use low-end systems to expand the low-end market while, at the same time, we need more competent systems (see items 1 and 2 above) to push the envelope.

Tom Tolles is the President and Co-founder of hOuse Of mOves motion capture studios. hOuse Of mOves is now in its 7th year of business and has grown to be one of largest providers of motion capture studio services. hOuse Of mOves's credits include motion capture work on feature films such as "Titanic", "Mortal Kombat: Annihilation", "Practical Magic", commercials for Coca-Cola, Mazda, Simmons Mattresses, TV series including Johnny Quest, and video games such as NBA 2K, NFL 2K, Madden 99, Tekken3, Knockout Kings, Parasitve Eve, and Tomorrow Never Dies. Although Tom has been working directly with motion capture for seven years, he has also been working in the CG industry for 17 years. In the mid-to-late 80s, Tom was the co-founder of FTI West, an animation production company that was located in San Francisco. Prior to forming House of Moves, Tom was Vice President, Sales & Marketing, Viewpoint DataLabs. Tom is a graduate of the Anderson School of Management ('91) and also holds a BSME/MSME from Stanford ('81/'82).

## 6 Barb Helfer

Barbara Helfer is a Senior Graphics Researcher at ACCAD, a center for graduate study in computer animation and movement housed in the College of Arts at the Ohio State University. Her teaching focus is in motion capture, digital video, and compositing aspects of computer graphics with video. She has worked as a course reviewer since 1993 as well as being the SIGGRAPH Courses Chair in 1997 and 1999.

## 7 Suba Varadarajan

Suba Varadarajan is a Graphics Research Specialist at ACCAD at the Ohio State University (OSU) and is in charge of the College of the Arts motion capture facility. She received her M.S. in Computer Science from OSU and is currently working on a Ph.D. Her area of research is in 3D model-based image processing of human figures and is interested in motion editing and re-use.

# Interactive Stories: Real Systems, Three Solutions

Noah Wardrip-Fruin<sup>1</sup>  
Brown University  
Moderator / Co-  
Organizer

Andrew Stern<sup>2</sup>  
Interactivestory.net  
Panelist / Co-  
Organizer

Peter Molyneux<sup>3</sup>  
Lionhead Studios  
Panelist

Michael Mateas<sup>4</sup>  
Carnegie Mellon  
University  
Panelist

Bernard Yee<sup>5</sup>  
En-Tranz  
Entertainment  
Panelist

## 1 Introduction

While pure theory can be interesting, this panel starts where the rubber meets the road for interactive stories – real systems, and the practice-oriented insights of creators. Expect to hear concrete examples and solutions to hard problems, ranging from the technical to the artistic. Of course, there's also a lot of controversy about what interactive stories are, and how to best make them, which this panel doesn't plan to ignore. It includes people working on major systems that represent three different approaches, and puts them together with a moderator that knows the hot buttons (as well as the points of agreement). The three approaches can be called: narrative game, interactive drama, and massively multiplayer storytelling.

Peter Molyneux, perhaps best known as the designer of the recent hit *Black and White*, is a master of the narrative game. This practice, growing out of computer game design, was pushed to new levels by the experience arc of *Black and White*. Now it promises to take yet another leap with Molyneux's upcoming project, which weds hero narrative with deep world simulation. This practice, embedded in a tradition of great gameplay, can sometimes find meaningful story its greatest challenge - but it's one Molyneux has been consistently able to meet.

Andrew Stern, when designing for the *Petz* series, created some of the most compelling virtual characters ever seen on a personal computer screen. This focus on character, and on engaging character interaction, is one of two major components of interactive drama - the other being a method for managing this interaction so that it results in a compelling story, to which collaborator Michael Mateas of *Terminal Time* fame brings to bear his experience as creator of next-generation AI-based artworks. While getting the character-plot combination right has proven quite elusive, pulling it off could open the door to the often-imagined world of interactive cinema.

Bernard Yee is a leading figure in the field of massively-multiplayer (MMP) gaming, where he has been involved with defining projects such as Sony's *Everquest*. In MMP worlds authors, rather than storytelling in a traditional sense, provide affordances and environments for players to tell stories collaboratively, through their actions. This *abdication of authorship* takes simulation-centered story well beyond the narrative game, but has so far only attracted a hard-core role-playing game audience. In the next two years, that may change.

This panel will be the first time leading representatives of these three approaches have appeared together at SIGGRAPH.

## 2 Peter Molyneux

Stories. Why sit back and enjoy them when you can actually MAKE them? Games are the only conduit by which you can be both a character in a story and the storyteller or narrator. In *Black*

& *White* we designed the game so that the player could do whatever he or she wanted. The big story of warring gods and planetary domination was always there, but how and when it played out was down to whoever had their hand on the mouse. Of course, we had to steer the narrative, but the main choices were left open. Do you betray and attack your ally? Do you opt out of conflict or do you side with him and fight your common foe? We catered for each eventuality. The system worked, but what we learnt was that although people expect to have their choices pre-guessed in games they feel far more fulfilled when their choices appear to have an impact on every aspect, no matter how tiny, of their game lives, and the world in which they exist.

We learnt this playing *Black & White*, but it was too late to implement it fully in that game. But in *Project Ego* we're making sure that the player does feel that everything he or she does really makes a difference. The gap between the game world and the big, overall story is being closed and as a by-product we're finding that it doesn't just take massive battles and toppling empires to make a player feel immersed in a game. The subtle touches work – people in your village mentioning that life is easier since the overbearing village elder joined his lord and was killed in the castle siege you started. And with *Ego* we can also show the changes on the detailed dynamic landscape.

So what we've learnt is that story isn't all about big, seamless plotting. It's about what we show the player as the game unfolds. Tiny details work as well as big events. Just as history is about daily life as well as presidents, kings and emperors, the stories we're working on are rich, deep and satisfying. Of course, they're also epic, but epic tales aren't the problem. It's the tapestry of life we weave around them that we've realized is so important. God may dwell in the details, but the future of god games definitely dwells in the details.

## 2.1 Molyneux Bio

Peter Molyneux is one of the best-known names and most articulate and eloquent speakers in the international world of computer games. He co-founded Bullfrog Productions in 1987 and created a new genre of computer games, "the god game" with the release of *Populous*. Since then Peter has been responsible for a string of massive selling games including *Powermonger*, *Theme Park*, *Magic Carpet* and *Dungeon Keeper*. Cumulative sales of his games are now approaching the ten million mark worldwide. In 1997 Peter left Bullfrog Productions to form a new games development company Lionhead Studios. The company's first game *Black & White* was released to wide spread critical acclaim in April 2001 and sales currently top the one and a half million mark. He has spoken at the American Museum of the Moving Image, the British Film Institute, BAFTA, GDC, GDCE, E3, ICA (London), the Tate Gallery and the Dortmund Museum of History and Culture.

## 3 Michael Mateas and Andrew Stern

Façade is an artificial intelligence-based art/research experiment in electronic narrative – an attempt to move beyond traditional

<sup>1</sup> nwf@brown.edu

<sup>2</sup> andrew@interactivestory.net

<sup>3</sup> pmolyneux@lionhead.com

<sup>4</sup> michaelm@cs.cmu.edu

<sup>5</sup> bernie.yee@mail.com

branching or hyper-linked narrative to create a fully-realized, one-act interactive drama. Integrating an interdisciplinary set of artistic practices and artificial intelligence technologies, we are completing a three year collaboration to engineer a novel architecture for supporting emotional, interactive character behavior and drama-managed plot. Within this architecture we are building a dramatically interesting, real-time 3D virtual world inhabited by computer-controlled characters, in which the player experiences a story from a first-person perspective. *Façade* will be publicly released as a free download in early 2003.

You, the player, using your own name and gender, play the character of a longtime friend of Grace and Trip, an attractive and materially successful couple in their early thirties. During an evening get-together that quickly turns ugly, you become entangled in the high-conflict dissolution of Grace and Trip's marriage. No one is safe as the accusations fly, sides are taken and irreversible decisions are forced to be made. By the end of this intense one-act play you will have changed the course of Grace and Trip's lives – motivating you to re-play the drama to find out how your interaction could make things turn out differently the next time.

This work is unlike any hypertext narrative or "interactive fiction" to date in that the computer characters actively perform the story without waiting for you to click on a link or enter a command. Interaction is seamless as you converse in natural language and move and gesture freely within the first-person 3D world of Grace and Trip's apartment. AI controls Grace and Trip's personality and behavior, including emotive facial expressions, spoken voice and full-body animation. Furthermore, the AI intelligently chooses the next story "beat" based on your moment-by-moment interaction, what story beats have happened so far, and the need to satisfy an overall dramatic arc. An innovative text parser allows the system to avoid the "I don't understand" response all too common in text-adventure interactive fiction.

### 3.1 Mateas Bio

Michael Mateas' work explores the intersection between art and artificial intelligence, forging a new art practice and research discipline called Expressive AI. He is currently at Carnegie Mellon, where he is an adjunct faculty member in the Entertainment Technology Center, a Research Fellow in the art department's Studio for Creative Inquiry and a Ph.D. student in Computer Science. Michael's AI-based artwork includes *Terminal Time*, a mass audience, interactive, story generation machine which constructs ideologically-biased documentary histories in response to audience feedback. Michael has presented papers and exhibited artwork internationally including SIGGRAPH, the New York Digital Salon, AAAI, the Carnegie Museum, the Warhol Museum, and Sonic Circuits (the Walker Museum entry). Michael received his BS in Engineering Physics from the University of the Pacific and his MS in Computer Science (emphasis in Human-Computer Interaction) from Portland State University.

### 3.2 Stern Bio

Andrew Stern is a designer and programmer of the interactive characters *Dogz*, *Catz* and *Babyz* from PF.Magic in San Francisco. Along with his fellow creators Adam Frank, Ben Resner and Rob Fulop, he has presented these projects at a variety of conferences including Digital Arts and Culture, SIGGRAPH 2000 Art Gallery, AAAI Narrative Intelligence Symposium, Autonomous Agents, Intelligent User Interfaces and GDC. *Babyz* won a Silver Invision 2000 award for Best Overall Design for CDROM. *Catz* received a Design Distinction in the first annual *I.D. Magazine* Interactive Media Review, and along with *Dogz* and *Babyz* was part of the American Museum of Moving Image's

Computer Space exhibit in New York. The projects have been written about in *The New York Times*, *Time Magazine*, *Wired* and *AI Magazine*. He holds a B.S. in Computer Engineering from Carnegie Mellon University and a Masters degree in Computer Science from the University of Southern California.

## 4 Bernard Yee

My thoughts on interactive story come from several sources, not just my experience with massively multiplayer environments (MMPs). The important thing about MMPs, in this context, is that they provide an opportunity to watch emergent narrative and gameplay grow from the interactions of many users. This happens even in static systems, as long as the systems offer consequences for player action. Beyond this, I think the perspectives most central to my thoughts on interactive story are the simulation-based approach to narrative design and the related phrase coined by Doug Church and a few other developers, the *abdication of authorship*. Designers have been too enamored of their own cleverness – making players figure out their interpretation of the story (especially in an RPG context) rather than creating systems that enable the player to express themselves via their game actions. Scripted quests are not really interactive at all, since they play "guess the designer intent" games with the player. Simulations seem to be the way to, instead, create game systems that allow multiple means of solving problems. Simulations can create a loose narrative, a mini-narrative, and/or allow the player to superimpose his own narrative on top of game events. An example of this from RL is the last few seconds of a tight NBA game – do you throw it inside and post up your big man, or kick it out for the open jump shot? That mini-narrative arc is a familiar one, creating tension and multiple ways to resolution.

### 4.1 Yee Bio

Bernard Yee brings a deep knowledge of online gaming business and design, ranging from business models to development to maintenance. As an editor and analyst, he covered the emerging online gaming industry starting in 1994, authoring the first market report on online gaming for Jupiter Communications in 1996. Before joining En-Tranz, Bernard was Director of Programming at Sony Online Entertainment, publishers of *EverQuest* and the upcoming MMP games, *Planetside*, *Star Wars Galaxies*, and *Sovereign*. His prior work also includes being Director of Creative Development at Disney Online/ABC Online, and Director of Product Development at ABC Interactive. Bernard attended Columbia University and Duke University Law School, and is an expert in entertainment and corporate law.

## 5 Wardrip-Fruin Bio

Noah Wardrip-Fruin both writes interactive stories and writes about them. His current nonfiction work includes being the lead editor of *The New Media Reader* (with Nick Montfort) and of *First Person: New Media as Story, Performance, and Game* (with Pat Harrigan), both of which are forthcoming from MIT Press. His current fiction work includes being the lead writer/designer of *The Impermanence Agent*, which tells a story, monitors the user's web browsing, and uses browsed materials to customize its story out of existence. The *Agent* has been featured in the SIGGRAPH Art Gallery, was shortlisted for the first Electronic Literature Organization fiction prize, and was part of the Guggenheim Museum New York's 2001 "Brave New Word" program. Wardrip-Fruin spent seven years as a research scientist and artist and residence at New York University, was a founding faculty member of the University of Baltimore's School of Information Arts and Technologies (SIAT), and is currently a Creative Writing Fellow at Brown University.

# Games: The Dominant Medium of the Future

## *Organizer*

Bob Nicoll  
Graphics Training Manager  
Electronic Arts  
[bnicoll@ea.com](mailto:bnicoll@ea.com)  
650.628.1500

## *Panelists*

Glenn Entis  
Patrick Gilmore  
J.C. Herz  
Alex Pham  
Will Wright

## *Moderator*

Ken Perlin

Driven by trends in silicon and software, computer gaming is the medium that will define the recreational and cultural experience of the twenty-first century in the way that motion pictures and their offspring television, defined the recreation of the twentieth. Or so we assert! This panel will debate the truth of the statement that gaming is the dominant medium of the future with believers, skeptics, and outside referees.

### *Glenn Entis*

**Position:** While we will all have fun arguing over the meaning of “dominant”, it is harder to argue about the games industry’s explosive growth relative to any other form of entertainment. The average gamer in the U.S. is now 28 years old -- and the sweet spot continues to migrate upwards as people who grew up playing games take the habit into adulthood. Also promising is the recent arrival of games that appeal to women such as *The Sims* and *Harry Potter*. Best of all from a demographic standpoint is the industry’s attraction of younger gamers. Generation Y, the mid-point of which is about to turn 12, is 15 percent bigger than generation X and totally immersed in the game culture.

**Bio:** Glenn Entis, VP, Electronic Arts Worldwide Studios focuses on the question “how does EA share knowledge and connect the right people” across its studios worldwide, and also leads EA Canada’s Games group in their mission to create new character based games. Prior to that, Glenn was the first CEO of DreamWorks Interactive in Los Angeles (now EALA), where he worked closely with Steven Spielberg and Jeffrey Katzenberg to bring their vision into games. Glenn oversaw the development of more than 15 titles at DWI, including the award-winning *Medal Of Honor* series and last year’s critically acclaimed PC game, *Clive Barker’s Undying*.

Glenn was a co-founder of the Oscar winning studio Pacific Data Images and for his work there received a Technical Achievement Award in 1998 from the Academy of Motion Picture Arts and Sciences. He and his wife live in Vancouver with their two pre-teen daughters.

### *Patrick Gilmore*

**Position:** The gaming industry's worst enemy is the idea, held by some self-congratulatory geeks, that games will be the medium to define the cultural experience of the 21st century. How can games influence a culture when most of them only imitate the drama of motion pictures, the action of spectator sports, or the adrenaline of extreme recreation? Comparisons of the gross

annual revenue of the games industry to first run theatrical motion pictures are nothing more than spinning the numbers. The real evidence lies in the culture itself, which largely continues to dismiss games as antisocial, overly violent and nerdy. But, hey, if you believe that narrow demographic of dateless, young adult men to be the cultural pioneers of the coming century, relax, kick back on your self-awarded laurels, and get ready for a hundred years of disappointment.

**Bio:** After producing video and computer games for nearly 10 years, Patrick Gilmore left the interactive industry to direct a major animated motion picture. In the film industry, the phrase, “it looks like a video game” is used disparagingly, but Gilmore has held on to an idealistic view of the potential of games. It’s when he hears that games have already arrived as a dominant cultural influence that Gilmore really bristles. It will be decades before games offer the depth of experience to the breadth of audience offered by film.

### *J.C. Herz*

**Position:** “Will games be the dominant medium of the 21st century” is one of those questions that’s fun to ask, and fight about in the attempt to answer, but doesn’t really lead anywhere useful. Putting media mano a mano is like putting comic book characters or superheroes mano a mano: Spiderman vs. the Silver Surfer! James Bond vs. Shaft! Games vs. Movies! When in fact, games and movies don’t really compete, any more than Shaft and James Bond compete. They do different things, in different places, at different times, often in tandem (i.e. *Star Wars*). I think the big question (and a much more difficult question) is how games and movies are fundamentally different, and what those distinct vocabularies and grammars do to the relationship between media, both in the marketplace and in popular culture.

**Bio:** J.C. Herz is the CEO of Joystick Nation Inc., a consultancy that applies the principles of complex systems to the design of products, services, and brands. Drawing from an understanding of ecology, online social dynamics, computer games, and information theory, Joystick Nation's focus is human-human interaction design, and systems that leverage the intrinsic characteristics of networked communication. Clients include multinational corporations, high-tech start-ups, nonprofit organizations, and the military.

Prior to founding Joystick Nation, J.C. was the author of two books, *Surfing On the Internet: A Nethead's Adventures Online* (Little Brown, 1994), a pop anthropology of life in cyberspace and *Joystick Nation: How Videogames Ate Our Quarters, Won Our Hearts, and Rewired Our Minds* (Little Brown 1997, between



editions in North America, British edition available at amazon.co.uk), a history of videogames, which traces the cultural and technological evolution of the first medium that was born digital, and how it shaped the minds of a generation weaned on Atari.

Continuing her investigation into the grammar and syntax of game design, J.C. wrote 100 essays about computer game design for the New York Times (her "Game Theory" column, which ran weekly through February 10, 2000, is archived on the Times' site, [www.nytimes.com/library/tech/reference/indexgametheory.html](http://www.nytimes.com/library/tech/reference/indexgametheory.html)). She has conducted workshops on game design and learning, and has spoken at technology, design, and business conferences including Technology Entertainment Design (TED) in Monterey, Game Developers' Conference, SIGGRAPH, E3, and the World Economic Forum. J.C. sits on the National Research Council's Committee on Creativity and Information Technology.

#### Alex Pham

**Position:** You're all delusional. Just because some guys in a basement have the technology doesn't mean the public wants it. The consumer wants a simple, focused experience. Until you can define that experience, you can neither design nor implement it, let alone sell it. And until you can sell it to a mass market, you can't "influence culture." Is a bunch of nerds playing Doom after hours really a "defining cultural moment" in the history of humankind? After billions of dollars spent developing gee-whiz technologies and hopelessly-confusing interaction models (Majestic as Trespasser redux?), we've learned that people pretty much like to click on pictures of buildings and people. Why even bother with virtual reality when the real thing is so much better? And why hassle with interactivity when passive viewing yields a more intimate experience? Football coach Bill Parcells once said, "I guess that's why they play the games," but he was talking about the real thing, not the John Madden simulation. Truth is stranger than cut-scene fiction, plus the sound effects are better.

**Bio:** Alex Pham has been a journalist for 15 years. Prior to coming to the Los Angeles Times in 2000, she worked for USA Today, The Washington Post, The Oregonian, and most recently The Boston Globe, where she wrote a weekly technology column. Alex currently covers the interactive entertainment and semiconductor industries for the Times. She has a Bachelor of Arts degree in International Relations from Johns Hopkins University and a Masters degree in public policy from the University of Texas at Austin.

#### Will Wright

**Position:** From a designer's point of view interactive media has (by far) the largest design space of any media type. The interactive design palette includes elements of theatre, books, movies, toys, art, architecture, psychology and many other fields. We've only begun to scratch the surface of this new landscape.

**Bio:** Will Wright is the creator of both the *SimCity* and *The Sims* franchise. *SimCity* was released in 1989, and within a few months became a hit. The latest incarnation and definitive version of *SimCity*, *SimCity 3000 Unlimited*, has also continued in the tradition. Wright's game *The Sims*, puts players in charge of the lives of a neighborhood of simulated people. Released in February

of 2000, this title has become a cultural phenomenon and sold over 6 million copies worldwide and has become the best selling PC game of all time. *The Sims* has inspired four expansion packs; *Livin' Large*, *House Party*, *Hot Date* and due out this March, *Vacation*. Combined sales for *The Sims*' franchise are approaching total 12 million units life-to-date. Next up for Wright is *The Sims Online*™. Scheduled for release in the second half of 2002, *The Sims Online* will enable you to take your Sims to an online world where you get to be yourself or whoever you want to be.

#### Ken Perlin - Moderator

**Bio:** Ken Perlin is a Professor in the Department of Computer Science and the director of the Media Research Laboratory at the Courant Institute of Mathematical Sciences of New York University. He is also the director of the NYU Center of Advanced Technology, sponsored by the New York State Science and Technology Foundation. He completed his Ph.D. in 1986 from the New York University Department of Computer Science. His dissertation received the Janet Fabri award for outstanding Doctoral Dissertation. He received his B.A. in theoretical mathematics at Harvard University in 1979. His research interests include graphics, animation, and multimedia. In 1991 he was a recipient of a Presidential Young Investigator Award from the National Science Foundation. In 1997 he was a recipient of a Technical Achievement Award from the Academy of Motion Picture Arts and Sciences for his noise and turbulence procedural texturing techniques, which are widely used in feature films and television.

Dr. Perlin was Head of Software Development at R/GREENBERG Associates in New York, NY from 1984 through 1987. Prior to that, from 1979 to 1984, he was the System Architect for computer generated animation at Mathematical Applications Group, Inc., Elmsford, NY. TRON was the first movie for which his name got onto the credits. He has served on the Board of Directors of the New York chapter of ACM/SIGGRAPH, has been a member of ACM and ACM SIGGRAPH, and is currently a member of the board of Directors of the New York Software Industry Association.

#### Bob Nicoll – Organizer

**Bio:** Bob Nicoll is a graphics trainer at Electronic Arts, and worldwide graphics manager of EA University. He has been both a CG production artist and trainer for the past twenty-five years. His experience spans university professorships, 3D CG artist, Art Director, Designer, Producer, Technical Support Specialist, and, most importantly, a teacher at Electronic Arts. His current focus encompasses technical and aesthetic design skills, from cg animation techniques to life drawing and improvisation classes.

# sigKIDS – Introduction

Education is an unlimited process. It takes place in many settings beyond formal classrooms: homes, streets, libraries, clubs, malls, and museums. This year, as it did in New Orleans for SIGGRAPH 2000, sigKIDS focused on involvement with the local community and reaching out to audiences that are not usually associated with computer graphics and interactive techniques.

SIGGRAPH 2002 sigKIDS offered two types of activities at the San Antonio Childrens Museum, a unique learning center designed to provide interactive and stimulating educational experiences that can be shared by both parents and children:

1. Camp sigKIDS for children 4-8 years old. The educational “camp” gave younger children their first opportunity to interact with technology-based learning materials. Registration for the camp was free and open to both museum and conference attendees. With SIGGRAPH 2002’s assistance, the museum received 12 computers from the San Antonio Courier Express and a variety of educational and other software. These permanent donations enabled the museum to develop its first hands-on lab experiences. The children of south Texas will be using the new resources for several years to come.
2. A variety of interactive educational projects, similar to those displayed in SIGGRAPH 2002 Emerging Technologies. These projects, submitted by specialists from the US, Canada, and Japan, celebrate learning in the sciences, cultures, and the visual and performing arts.

Marc J. Barr, Chair

# sigKIDS – Committee

## Chair

Marc J. Barr  
Middle Tennessee State University

## Committee

Guanping Zheng  
Middle Tennessee State University

Christine Sinick  
San Antonio Children's Museum

Alethea Hopkins  
San Antonio Children's Museum

Deborah Chew  
Birney Elementary School

Brent Walters  
DigitalRenderings.com

# Interactive animation as an educational tool in "Winter Dreams"

Daria Tsoupikova  
Computer Graphics  
Syracuse University  
datsoupi@hotmail.com



## Abstract

Being interested in Interactive Animation, Illustration, Multimedia, and the creative use of programming languages, I explored these areas to produce an educational game for preschool children. The following four chapters present the concept and development of an experimental interactive game titled "Winter Dreams".

## 1 Introduction

"Winter Dreams" is an experimental interactive multimedia project. This game was developed for kids age five to seven, and teaches the following skills: object recognition, picture identification and matching, early math and problem solving, and creativity. The game consists of chain-style interfaces that require the child to solve simple game tasks in order to earn rewards and move to the next interface. Simple text explanations at the bottom of the window, symbols, and animations help him to understand the task.

## 2 Concept and the story

"Winter Dreams" is a narrative-based game about the winter hibernation of a bear family. It consists of two major interfaces and five dream scenes describing Mother Bear in the act of putting her five children to sleep. The first interface introduces the project's theme to the user. The second allows the user to explore the five winter dreams and the stories which mother bear tells her children to make them fall asleep. Each of the five dream scenes has different interactive elements, surprises and animations depending on the theme of the dream. The user can interact with them through mouse controlled actions, sound and visual surprises and mouse-manipulated objects. Interaction adds further developments to each dream story. The audio side consists of the dream-narration recordings and of different sound effects, which range from loud yawns to quiet music. The user can listen to the stories by clicking on each of the bears to find out what dream a bear sees while asleep. After a dream has been viewed, and the little bear is asleep, the user is returned to the main scene. The main purpose of this game is to put all the little bears to sleep.

## 3 Interactivity as a part of the concept

I believe interactive animation is a very powerful educational method in K-12 pedagogy, and consider the game interface as one of the best ways to reach children. It provides the child with a great source for exploration, and discovery, by balancing interactive control, with different visual surprises. All these elements contribute to the game's learning environment and help the user to navigate between the different parts. Some of the interfaces have a layered control structure, where the interface can be changed by the user. Other interfaces seem to be fixed at first, but later, while the animation unfolds, provide the user with some additional choices to interact with.

There are two different learning opportunities for the user in this project. The first mode provides consistent game rules and tells the user about where to go next, what to do, and what to anticipate. Some mouse clicks will cause predictable results. Users can solve tasks in each interface if they follow the rules. They can make decisions on how they want to play, and they have the control over the story.

The second learning mode consists of visual surprises, and unexpected, random results. This part provides the user with a great opportunity for exploration and creativity. The user can change the game interface and play with the game elements. But even though the game is controlled by the users and gives them the opportunity to change the order of the dreams, it also maintains the traditional final "Happy ending" by having, by the end of the game, all the little bears peacefully asleep "until next spring."

## 4 Development and Influences

I wanted to combine my interests and explorations in animation, illustration, sound/image interrelation and interactivity with the creative use of a programming language. The production process included several development stages, from concept design and storyboard development, to programming and Web production. First I drew the images using watercolor on paper. Then I scanned and redrew the images entirely, using the Flash drawing and painting tools. Although I used a great variety of colors in my illustra

tions, I tried to work within the frames of color theory requirements.

"Winter Dreams" is based on the "Sleepy bears" story by Mem Fox, and the Russian fairy tale "Three Bears". I was also influenced by Ivan Bilibin, a great Russian artist who lived and worked in Moscow in the beginning of the last century. He extended the language of lines and colors, and defined new modes of Illustration. I found myself always looking back to his masterpieces, especially these days when most online children animations use rather restricted color palettes. His investigations of new shapes and lines, and his research in color values were a great source of inspiration for my project.

## References

FOX M. 1999. *Sleepy Bears*. Harcourt Brace & Company.

BILIBIN I. 1899-1900. *Vassilisa the Beautiful*. GOZNAK, Moscow 1976.

# Gaming as an Educational Tool: Internet Scavenger Hunt

Karen Monahan\*  
Parsons School of Design

Can games be educational as well as fun?

Educators and children's media developers have long been trying to answer this question. Much of the software produced for kids and marketed as educational is comprised of the type of games that Seymour Papert has so adeptly defined as "edutainment".<sup>1</sup> Many of these games are drill-type memorization exercises where the game activity has little or no relation to the "educational" component. Mainstream educational software might be fun, even engaging, but very little of it challenges the user intellectually or encourages open-ended discovery.

*Go Wild Scavenger Hunt* is a web-based educational game for children ages 8-12 that encourages and teaches kids how to explore the Internet through an investigative virtual scavenger hunt. The game technology is transparent, allowing the user, through hands-on experience, to develop an understanding of exactly how computers work. At the same time, the user is introduced to an unfolding world of animal science. Game play is organized to foster self-initiated learning as well as cooperation and communication among participants.

The goal for players is to win as a team in a weekly competition against two other teams to discover animal taxonomy. The three teams are Flying Creatures, Land Inhabitants, and Water Dwellers. Players search the Internet to find pictures of animals within an assigned animal class, such as "Crustaceans". The animals selected by each team must inhabit that team's environment (air, land, or water). Tools to assist with Internet searching are provided, as well as links to helpful children's search engines.

This game is all about hands-on learning. Since there are multiple right answers for each team every week, users are encouraged to experiment and develop a variety of strategies to gather information on the web. Moreover, players can comment on the resources and facts they've discovered, share this knowledge with teammates, and check and challenge their opponents' results, effectively teaching each other. Through this exchange of information, not only does the individual player build her own knowledge about animals and computers, but she also gains experience in helping both her teammates and the larger community to do the same.

This online scavenger hunt demonstrates just one possible solution to the problem of designing applications that emphasize and enhance the educational function of technology.

## References

1. PAPER, S. 1998. Does Easy Do It? Children, Games, and Learning. *Game Developer*, 88.



Figure 1: Game introduction and instructions.



Figure 2: Sample of a team interface.

\*email: monahank@newschool.edu



# FORM

Hilary J. Wright\*  
Nancy Hyland

## 1 Overview

*Play is the highest expression of human development in childhood for it alone is the free expression of what is in a child's soul.*

- Fredrick Froebel

FORM is an educational software prototype designed to encourage children's creative play with fundamental geometric shapes. The open-ended program serves as a creative tool to help form a deep and lasting understanding of the beauty, elegance and underlying unity of math, science, design and nature.

## 2 Description

Through play with simple geometric shapes, children are introduced to: design and aesthetics (color, symmetry, pattern); geometry (shape, proportion, harmony) and physics (volume, gravity, movement). The hexagon grid interface invites players to create static and animated designs by changing the size, weight, color, position, gravity, animation and sound of circles, squares and triangles. Audio, animated clips and graphic icons help pre-readers to navigate through the program independently.

Off-computer activities include: a 3D cube, sphere and cylinder; 2D colored shapes; a grid play area; and an activity booklet with suggestions for adults to help children get the most out of FORM. One activity includes observing, gathering and comparing objects from nature such as shells, leaves and stones. In another, different color combinations are presented along with questions about how colors make one *feel*, encouraging discussion of the relationship between design and emotion.

## 3 Goals

We hope FORM will help children:

- improve spatial, symbolic and creative thinking skills
- develop strong visual communication skills
- explore relationships between design and emotion
- recognize their unique creative abilities
- perceive the unity of design, math, science and nature

## 4 Froebel and Play

In developing FORM, we have been profoundly influenced by Fredrick Froebel (1782-1852), German educator and inventor of kindergarten, who believed that the best way to introduce children to the beauty and interconnectivity of geometry, design and nature was through *play*. "The mind grows by self revelation. In play the child ascertains what he can do, discovers his possibilities of will and thought by exerting his power spontaneously. In work he follows a task prescribed for him by another, and doesn't reveal his own proclivities and inclinations; but another's. In play he reveals his own original power." [Froebel 1887]

\*email: form\_software@hotmail.com

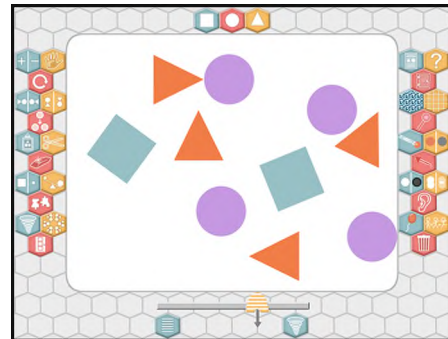


Figure 1: FORM prototype interface – *Wind*

Froebel developed a set of didactic materials, called "Gifts and Occupations", to introduce children to the fundamental shapes found in the natural world and, ultimately, to their own creative selves. He believed that each "Gift had a symbolic meaning, bringing forth from the child an innate idea of the Absolute and the mathematical laws that govern the universe". [Varnelis 1998]

In *Inventing Kindergarten*, Norman Brosterman [1997] uses the work of modern artists and architects who attended early kindergartens – among them Le Corbusier, Paul Klee, Wassily Kandinsky and Frank Lloyd Wright – to show that kindergarten's "universal, perfect, alternative language of geometric form" directly influenced the development of modern art. We hope that like Froebel's Gifts, FORM can help children develop a playful relationship to math, science and design, realize their own creative power and invent entirely new ways of understanding the world.

## 5 Future Potential

By previewing FORM at sigKIDS we hope to learn from children what works and what does not. We will then apply this information toward the development of a fully functional program. We are currently seeking funding and/or partnerships with educational software developers and welcome input from designers, developers and educators. We invite everyone to visit the San Antonio Children's Museum to talk with us about FORM.

## 6 References

- BROSTERMAN, N. 1997. *Inventing Kindergarten*. New York: Henry N. Abrams. p.12
- FROEBEL, F. 1887. *Education of Man*. New York: D. Appleton & Co. Translated by W.N. Hailman, AM. p. vi. Editors preface
- FROEBEL, F. 1912. *Froebel's Chief Writings on Education*. London: E. Arnold. Translated by S.S. Fletcher, MA, PhD and J. Welton, MA. p. 50
- VARNELIS, K. 1998 *The Education of the Innocent Eye*. Journal of Architectural Education, May

# The Floating Words for Kids

Interactive installation for learning alphabet

Satoko Moroi, Tokyo Denki University  
moroi@ia.dendai.ac.jp

Shinji Sasada, Japan Electronics College  
sasada@jec.ac.jp

Ryoji Shibata, Tokyo Denki University  
shibata@ia.dendai.ac.jp

## 1. Abstract

"The Floating Words" is a new device for playing with words. This installation provides the viewer with a new feeling: If you speak to the microphone, your voice drips into a water pool as water drops, and will begin to float. You can stir or ladle the "letters" with a stick or ladle.

The installation of version\_1 was exhibited at SIGGRAPH2001 Art Gallery, and it was succeeded very much.

This time, we add new expression to new version: "The Floating Words for KIDS", and the expression is device to learn alphabet playing with the "The Floating Words". The expression is: If you ladle some letters with a ladle, they will turn to the characters showing word that has the alphabet in an initial.

This project is a proposal of using this installation for learning alphabet for kids.

These feeling: it seem like a magic are results from tricks of combination of voice recognition system, 3D magnetic sensor, simple computer graphics and real water and so on.

## 2. Magical tools: Introduction and its mechanism.

### 2-a. The microphone

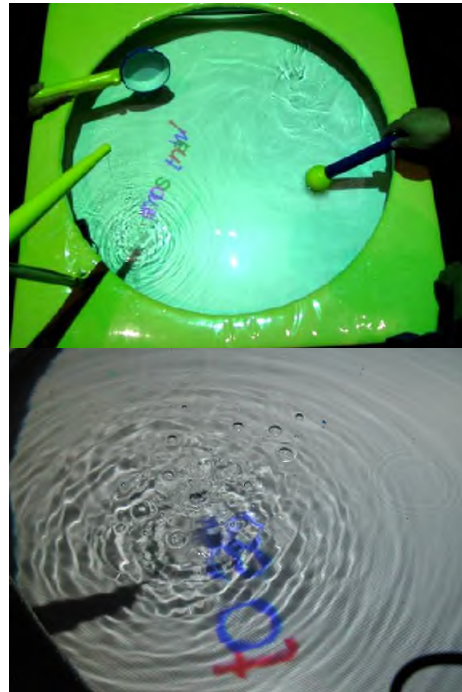
Through the microphone, your voice drip into a water pool as water drops, and they will begin to float as a letters.

When the letters appear into the water pool, It is just like magic.



The microphone sends voice to the computer, and the voice recognition system analyzes it, then make text file.

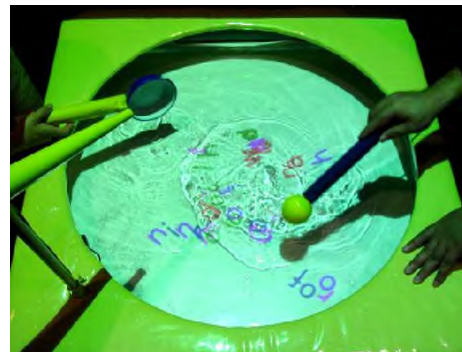
The microphone has a small pump under the pole. When someone speak to the microphone, it draw water from water pool because the microphone is connected to the pump controller. The pump controller sends electric power only when the microphone receives sound.



### 2-b. The Stirring bar

You can stir the "letters" with a stick or ladle. If you make flow of water, the letters float with the flow.

The stirring bar has a magnetic 3D digitizer inside, and it is sending motion data of it to the computer at real time.



### 2-c. The ladle

You can ladle some letters above to the water pool. If you ladle some letters, they will turn to the characters showing word that has the alphabet in an initial.

For example, "W" => whale, "E" => elephant, "T" => tulip.

And if you release it into the water pool, they return to a character, and float in the water again.

The ladle also has same mechanism as the stirring bar. And the sensor is sending motion data of it to the computer at real time. The computer will recognize constructing letters with a ladle, if the ladle rotates. And if it rotates once again, it will recognize having released the character.



### 2-d. Floating letters

Floating letters is computer graphics animation with OpenGL. The real-time animation is generated with reference to the text file from the speech recognition system (see 2-a) and motion data of ladle and bar (see 2-b,c), and the animation is projected on the bottom of a pool. Regarding water pool it has not special equipment but its bottom is white color. So it works not only as a water pool but also as a screen.

### 3. Consideration

In order to hallucinate letters as a more real object, it is important that the whole space is everyday as much as possible.

#### 3-a. Using real water

Water's weight, temperature, and a feeling of resistance feeds back to a hand. The feeling is important in order to raise a reality. In addition, since everyday water intervenes, we will forget that letters are digital virtual images by the computer.

#### 3-b. Not to touch special equipment

All the tools: microphone, a ladle, a stick, and a pool consist of simple forms and materials like a child's toy. The magic which appears from the tool with which we saw and got used gives big surprise by us.

#### 3-c. Space production

The production of the space which can be relaxed as much as possible is desirable. Since especially this time is taken for a child, it is important that the place has a bright atmosphere. Moreover, it is also important that accessories be made strongly so that it may be touched in comfort by the child.

And, as one more aims of this project, I think it is important that kids play with real water and share the time together with other participants for their mental growth of kids.

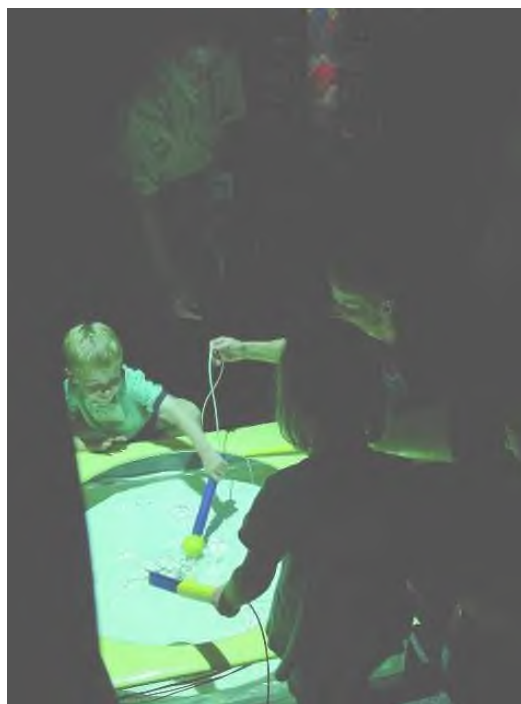
### 4. The cause of this project

Last year, At SIGGRAPH2001 Art Gallery, This installation version1 was exhibited as for adults. According to the last year of EAAC: Electronic Art and Animation Catalog, I explained like following. : "When your voice drips as water drops, when letters float in the pool, and can be stirred with a stick, you forget the everyday world, remembering instead the feeling from childhood when life was pleasant and time passed very slowly". This explanation was based on the grown-up viewpoint. Then our plan succeeded very much, and through last year's exhibition, we discovered that children were amused from another viewpoint. All of them was so excited. And they concentrated their attention on their work: speaking, stirring, and ladling. It seemed as if playing game.

We will install it anew for children as a device for learning the alphabet with concentration, enjoying oneself.

And we thank many people who gave various opinions to us and helped us on the occasion of this exhibition.

Further, it expects becoming exhibition significant for many children and teachers.





# Toys to Teach – Mathematics as a Collaborative Climbing Exercise

James Dai\*, Michael Wu\*, Jonathan Cohen\*, Troy Wu\* and Maria Klawe  
 E-GEMS Research  
 University of British Columbia  
 egems@cs.ubc.ca

\*these authors contributed equally to this work

## Abstract

Whereas traditional research on collaborative educational systems has primarily focused on how to define better modes of digital interaction, this approach is found lacking when applied to developing collaborative systems for elementary school aged children. It is creatively and collaboratively restrictive to filter the enthusiastic interactions of these excited 12-year-old children through progressively more complicated GUIs. Current E-GEMS research examines design factors of educational systems that recognize and facilitate the social context of the classroom and hopes to encourage, rather than restrict, peer-to-peer social discussion and interaction. By correlating observed interactions in the digital domain with those in the social domain, we hope to shed light on design factors of collaborative systems that can be an integral and exciting part of a child’s mathematical education. The vessel of our current research is the two-player collaborative mathematical exercise PrimeClimb, developed at E-GEMS. This paper describes the study conducted with PrimeClimb and documents the methodology of data capture and analysis used in the study – methods borrowed from ethnography, education research and sociology.

**Keywords:** *educational technology, collaborative learning, peer to peer discussion, computer support for collaborative learning*

## 1 Introduction

Since 1992, the E-GEMS lab at the University of British Columbia has been a playground for an interdisciplinary force of researchers in Education and Computer Science, professional game developers, school teachers and thousands of children. E-GEMS research focuses on using educational games as toys to teach children mathematical concepts [Klawe 1999]. Our results have demonstrated that games can be very effective in increasing both motivation and achievement in mathematical learning for girls and boys, playing apart and especially together [Inkpen 1995; McGrenere 1996].

The foundation of E-GEMS research includes commercially available games as well as professional games developed in the lab. Our games are iteratively designed to explore vital factors that promote player communication, accountability and thought about the underlying mathematical principles through the motivation of single and multi-player adventures. Our stories are weaved non-gender specifically and instead encourage the players to find their own identities through tackling the challenges in the games. Screenshots of past E-GEMS games are at the end.

PrimeClimb (Figure 1), in which a pair of children attempt to climb to the top of a mountain of numbers together, is the newest addition to the E-GEMS playground. In order to reach the top of each mountain, students must collaboratively avoid falling – a feat only accomplished by recognizing shared common factors between the numbers on which they stand. There is a tool – the magnifying glass – that can be used to shed light on the factors of any number. Magnifying glasses are hint-giving resources that are limited and shared amongst the two players. Our observation sessions were used to provide feedback in developing the current design factors of the magnifying glass.

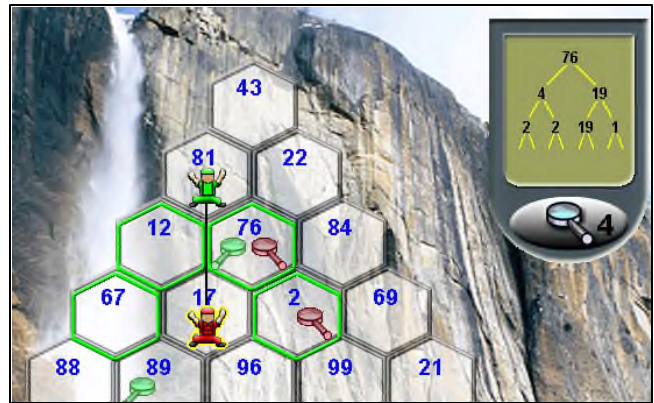


Figure 1 – PrimeClimb Screenshot

## 2 Research Setup

With PrimeClimb, we are primarily interested in analyzing design factors of hint-giving tools that promote collaboration in educational systems. Toward this end we have conducted a study in a local elementary school with two versions of the game to analyze the pedagogical effects of limiting a valuable shared resource - the number of magnifying glasses given to the players per level. The two versions used are in shown in Table 1.

	Tool Allocation	Sample Size
Version A	Limited	22 children
Version B	Relatively unrestricted	22 children

Table 1 - Sample Allocation

Version A allowed 6 possible uses of the magnifying glass per level – a number found to be rather restrictive in our prototype sessions. Version B had increasing allowances of the magnifying glass as the mountains became larger – the number given is approximately one magnifying glass per every 4 numbers on the

mountain – quite an abundant supply. Both versions of the game were played for 20 minutes by 11 pairs of grade 7 children (7 pairs of boys, 4 pairs of girls). In total 44 children were observed over the period of an entire week.

### 3 Methodology

The sessions of the study proceeded as follows: the pairs of children were first given a pre-questionnaire to assess their mathematic factorization skills and general attitudes toward collaboration in math and computer games. They then played PrimeClimb together for 20 minutes and the session ended with a post-questionnaire that reassessed their math skills and attitudes. Each play session was computer-logged and video-recorded. In order to correlate the digital events in the computer log with social discussion and interactions, an absolute clock was displayed on the game screen at all times. Each video captured a player's upper body along with their game screen; thus two tapes were produced from each session (Figure 2, the image has been blurred to preserve the subject's identity).

Captured audio transcripts have been successfully used to correlate social discussion and digital events in ethnography [Suchman 2000] and math education research [Pirie 1996]. We hope to use our captured video data to match social interactions with the PrimeClimb game events. By plotting both sets of events on the same absolute timeline (which is made easier through the displayed game clock on the bottom left corner of Figure 2), patterns between the digital and social worlds can be analyzed. The mixed video data captures the sessions more completely than audio transcripts and will allow us to supplement the social discussion with non-verbal physical interactions, such as the exchange of glances or gesture encouragement.

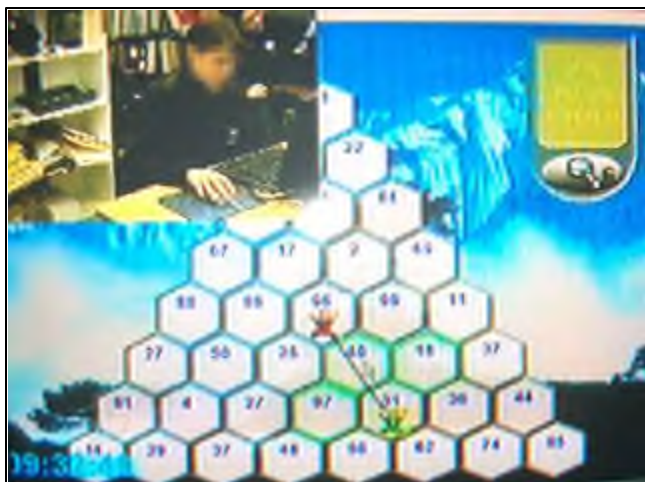


Figure 2 – Mixed Video Capture

### 4 Discussion

The data from the study sessions have been completely captured – including more than 20 hours of mixed video tape and all the pre and post questionnaires. The detailed analysis of the data is underway. Through our field observations we have already seen encouraging trends, three of which are elaborated upon below.

Firstly, we have seen that in the version of the game that has the more restrictive number of tool usages, the hint-giving tool played a much larger role in discussions about climbing strategy. For example, in this version of the game the players reminded each other of the presence of the tool and the number of tool uses

remaining much more frequently than the version of the game with the much less restrictive tool allocation.

Secondly, different categories of social interactions of children playing PrimeClimb are becoming clearer. As in past research from mathematics education, the categorization of peer-to-peer discussion in a mathematical context should take place first and foremost before the analysis of the pedagogical values of those categories can be carried out [Pirie 1998]. A few possible patterns of interactions in the context of PrimeClimb have emerged that include: a player explaining to the other player why a move is not allowed using mathematical language, a player explaining the reason for a failed move using mathematical knowledge and a player strategizing and logically planning a series of moves and explaining this strategy to their partner. An example of the latter is when a player states that she is on a prime number and encourages her partner to move as high as possible. Just as the categorization of peer-to-peer discussion in the classroom context has been subject to detailed analysis in mathematical education [Pirie 1998], we hope to apply this categorization and subsequent analysis of peer-to-peer discussion about math in the context of this computer supported collaborative environment.

Finally, we have noted that all the children had very little difficulty grasping the interface to the game – a result of multiple prototype sessions that have shaped the game interface. This means PrimeClimb could serve as a stable foundation upon which other research questions about collaboration can be launched. We noticed that girls and boys tended in general to employ different strategies in progressing through the game. Whereas some girls leaned more towards careful team-based planning and some boys were much more competitively collaborative and trigger-happy, the game accommodated all ranges of playing styles. It has been argued that designing games especially for girls is not a viable route because of the risk of marginalizing them – instead games should be non-gender specifically designed to allow children to find their own identity in the context of the challenge [Cassell 2002]. PrimeClimb seems to have succeeded in this respect.

### 5 Conclusion

The data collection part of the PrimeClimb study has been completed and the data analysis has begun with promising signs. Our ultimate vision is to build collaborative educational systems that fit seamlessly into and facilitate the social context of the classroom – this study has so far shown that ethnography and math education research have methodologies that can provide valuable insights.

### 6 Acknowledgements

The current E-GEMS research would not be possible without the vital support of many people. We would like to thank Kevin Atkins, Karen Wong and Jason Sze for their help with developing and prototype-testing PrimeClimb, and Wai-Chi Tsiang and Dr. Cristina Conati for their invaluable ideas and feedback. The most important part of E-GEMS are the children who are persistently enthusiastic toward our games, and toward our attempts at making math just a little bit more fun. Without the enduring support from the teachers - Val Day, Carolyn Dymond, Eileen Phillips and others – the children would not be there and the playground would be an empty one. Thank you all.

## References

- CASSELL, J. Genderizing HCI. In Press. 2002. Jacko, J. and Sears, A. (eds.), *The Handbook of Human-Computer Interaction*. Mahwah, NJ: Lawrence Erlbaum.
- INKPEN, K., BOOTH, K. S., KLAWE, M. AND UPITIS, R. 1995. Playing Together Beats Playing Apart, Especially for Girls. In *Proceedings of Computer Support for Collaborative Learning 1995*. Bloomington, Indiana.
- KLAWE, M. M. 1999. Computer Games, Education and Interfaces: The E-GEMS Project. <http://taz.cs.ubc.ca/egems/>
- KLAWE, M. M., WESTROM, M., DAVIDSON, K. AND SUPER, S. 1996. Phoenix Quest: lessons in developing an educational computer game for girls ... and boys. Presented at ICMTM96. <http://taz.cs.ubc.ca/egems/>
- MCGRENERE, J. 1996. Experimental Design: Educational Electronic Multi-Player Games – A Literature Review. Technical Report 96-12, the University of British Columbia.
- PIRIE, S. E. B. 1996. Classroom video recording: When, why and how does it offer a valuable data source for qualitative research? Paper presented at the 18<sup>th</sup> Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education. Panama City, Florida. ERIC ED 401 128, 17 pages.
- PIRIE, S. E. B. 1998. Working towards a design for qualitative research. In A. Teppo (Ed.), *Qualitative Research Methods in Mathematics Education*. Journal of Research in Mathematics Education Monograph Series, Reston, VA: National Council of Teachers of Mathematics, pp. 79-97.
- SEDIGHIAN, K. AND WESTROM, M. 1997. Direct Object Manipulation vs. Direct Concept Manipulation: Effect of Interface Style on Reflection and Domain Learning. *HCI 97: Premier European Conference on Human-Computer Interaction*. Bristol, UK.
- SUCHMAN, LUCY A AND TRIGG, RANDALL H. 2000. Understanding Practice: Video as a Medium for Reflection and Design. Taken from Baecker et al (2000). *Readings in Human-Computer Interactions: Toward the Year 2000*. pp. 233-240.



Figure 3 – Super Tangrams Screenshot [Sedighian 1997]

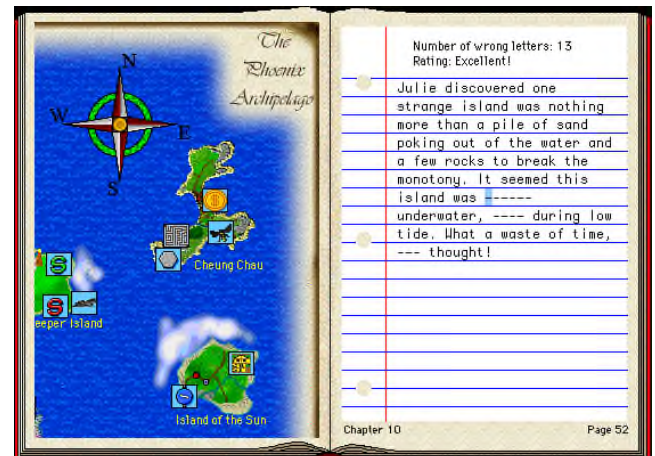


Figure 4 – Phoenix Quest Screenshot 1 [Klawe 1996]

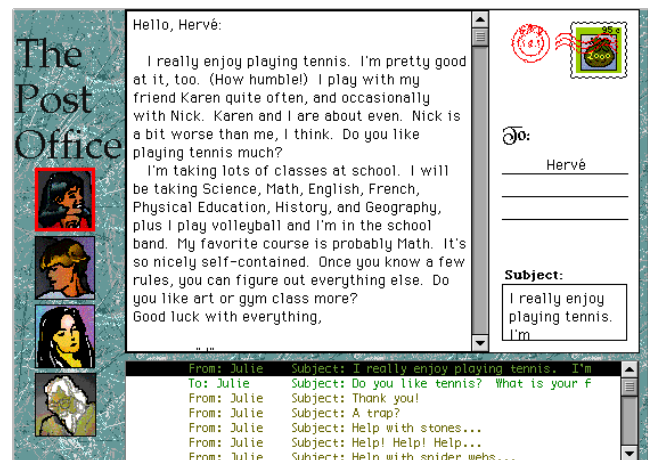


Figure 5 – Phoenix Quest Screenshot 2 [Klawe 1996]



# Jollee-Mail Playground

Yuichiro Kamata\*

Producer, Cyber Network Co., Ltd.

## 1. A Simulated “E-mail / World Wide Web” Environment for Children:

JOLLEE-MAIL PLAYGROUND is a simulated environment where children can experience the Internet communication in a more involving fashion.

The ever-evolving digital network is truly the infrastructure for the 21<sup>st</sup> century. To children, however, it is not a world so easily accessible. Also, it may be difficult for children, whose strength does not lie in their abstract (or logical) thinking, to understand the “virtual” world of the Internet.

This kiosk program is a playground for children where they can create their own unique e-mail with pictures and upload their own Web pages. This is meant for children to learn and experience as their having fun, but it is also exciting for grownups, who can rediscover the pleasure of exchanging e-mails in a more “creative” manner. Moreover, this playground environment isn’t close among the participating children, but it links the Web pages they create to the real Internet universe. We also believe we’re able to foresee how children will actually send e-mails in the near future (next year?) through this JOLLEE-MAIL PLAYGROUND.

**There are two kinds of sessions in this Internet-related program:**

### • E-mail = Mail Station

Which extends the fun of drawing/painting and Print Club (“puri-kura”) to the Internet. First you take a video footage of your own headshot and create a piece of “animated” painting over your own image, and then you send the artwork you’ve just finished by a simulated e-mail circuit as well as receive those done by other participants. You can touch up the received picture-mails or combine them with your own if you would like to pursue a more advanced level of modern art. The final result is printable.

### • Creating Web Page = Web Page Factory

Which enables you to create your own web page at ease. All you have to do is touch the screen, choosing your favorite template, draw or paint your artwork and write your message. And *VOILA!* Your own web page is built and ready! This “instant” web page can be uploaded into our server and seen by anybody from all over the world.

This project was originally developed as the main exhibition for Multimedia Zone, which is part of a Japanese science museum. Since then, it went through various experimentations with young participants. This current version now allows maximum of 12 participants to play it simultaneously (for the SIGGRAPH presentation, we might have to assemble a demo version with a less number of simultaneous participants).



Figure 1: Mail Station



Figure 2: Web Page Factory

## 2. Cartoon Character Represents Each Child Participant:

We use a network of personal computers as our simulated Internet environment. Whichever personal computer you may choose, all you need to do to participate in the Jollee-mail session is touch the monitor screen (for the SIGGRAPH presentation, this aspect depends on the availability of “touch screen” monitors).

A cartoon character leads you into the session, which you can play at your own playful pace. You will be guided into one of 12 Virtual Mail Rooms (or less for the SIGGRAPH presentation), each of which is also designated with a cartoon creature. In a sense, you will become one of those fanciful avatars and exchange your e-mail with the other characters. A child or not, you will enjoy the environment and can learn how to send a unique e-mail and create your own Web page in an easy and natural fashion.

\*e-mail: you@cyber-net.co.jp

We'll try to set up a multiple number of personal computers at other places than our booth in the Children's Museum so the experience can be made even more attractive. In addition, participants can log on from any location in the premise by using notebook PC connected by radio LAN, making the entire place a part of JOLLEE-MAIL PLAYGROUND.

Upon completion of the session, participants can stop by at our booth to pick up printouts of their own artwork created through the network as a souvenir to take home.

### 3. We Could Find a Future Possibility in Children's Playfulness

Children's playfulness might be a great source of new ideas and inspirations when we struggle to find out a better form of digital communication. Originally we created this program so that children can discover, at a museum, how enjoyable it is to communicate through the Internet. But we also feel there's a definite potential that this simulated environment of e-mail and website would give us many hints for a genuine improvement in the real world. In that sense, we believe it's truly meaningful to present at SigKIDS this JOLLEE-MAIL PLAYGROUND, which could easily turn into a common scene of children's home activity in the very near future. By simulating a future possibility in some concrete form, we feel we could contribute, however small it might be, to the progress and expansion of our global communication.

### 4. Development Team

This project was conceived and developed by Cyber Network who has been working for years on development of exhibition kiosk programs and various multimedia software, using digital technologies. We are really good at creating edutainment-type programs for museums. The company participated in SIGGRAPH in years 1994, 1995, 1997, 1998 and 1999. The basic development budget was provided by Niigata Prefectural Museum of Natural Science and Yokohama Science Center as our clients.

Cyber Network Co., Ltd.  
<http://www.cyber-net.co.jp/>

#### Contributor

Niigata Prefectural Museum of Natural Science  
<http://www.lalanet.gr.jp/nsm/>

Yokohama Science Center  
<http://www.city.yokohama.jp/yhspot/ysc/ysc/ysc.html>

Uchida Yoko Co., Ltd.  
<http://www.uchida.co.jp/>

Cyber Entertainment, Inc.  
 Nobuo Masuda  
 masuda@cyber-net.co.jp



Figure 3: Those "animal" addressees correspond to children actually playing at their booth. It helps those young people to understand better.

*"Yokohama Science Center"*



Photo 1: 12 children can participate simultaneously.  
*"Niigata Prefectural Museum of Natural Science"*



Photo 2: Younger children are really involved in this program.

# The Virtual Dig

Robert Dunn  
Arc Vertuel, Inc.  
rd1s+@andrew.cmu.edu

The Virtual Dig is an interactive adventure for young people and provides an educational experience in archaeology and historical architecture. The web site was produced for the Israel Museum, Jerusalem, in collaboration with museum archaeologists and can be seen online at the addresses: “<http://www.imj.org.il/arc-tel>” and “<http://www.arcv.org>”.

This 3D computer graphic reconstruction and walk-through is based on the permanent TEL exhibition at the Israel Museum. A “tel” is an archaeological mound. The physical recreated tel at the museum is a composite of five different historical periods each represented by architectural remains which reflect characteristics of specific periods.

Our Virtual Dig web site is a teaching tool which serves as a parallel educational resource for distance learning and for visitors to the museum who wish to understand the TEL exhibit in context and in depth. For example, online visitors interested in any period can study and compare architectural styles and the function of a place, view related artifacts and initiate an appropriate reenactment of an archaeological activity. Online visitors can reference digital photos of actual dig sites around the country which relate to the virtual reconstructed mound and see objects unearthed from a specific epoch that are part of the museum collection. Five historical periods are represented in chronological order: Neolithic, Canaanite, Israelite, Herodian and Muslim.

Source material for the 3D model included plans, perspective drawings, site photos, artist renderings, and detailed interviews and discussions with museum archaeologists, all collected during several visits to Jerusalem. Observations of children learning through play at the museum exhibit provided original data and insight for our design.

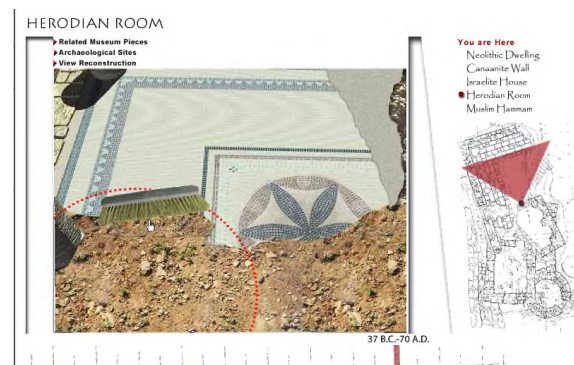
A walk-through of the 3D site orients the online visitor along a timeline of many centuries. Interactive engagement and graphics-rich settings communicate a sense of place and stimulate investigation.

3D CG reconstructions are used to represent the incompletely revealed and the time-worn, i.e. partial archaeological strata which exist simultaneously in the ancient landscape. Interactive animations are used to simulate the discovery processes of archaeology including digging and brushing, revealing jars and mosaic floors or demonstrating the purpose of a space. Each historical dig stratum has associated with it a digital photo gallery of artifacts from the museum’s collection. Some dig strata have links to complete 3D historical room reconstructions. These include the Herodian room and the Muslim hammam. A timeline and a site plan with elucidated viewpoints bracket the walk-through image frame.

This Virtual Dig web site for young people promotes a desire to visit the onsite permanent exhibit, sparks curiosity and reinforces memories of prior visits to the museum, provides a form of distance-learning for individuals seeking an opportunity



3D overview of the Virtual Dig.



Revealing the mosaic floor, Herodian room.

to experience the past, inspires play and exploration, exposes young learners to the processes of investigation associated with archaeology, and showcases the potential of virtual representation and simulation.

Design software used includes Cinema 4D, Flash, and PowerAnimator.

This unique educational collaborative project includes contributions from Dept. of Architecture students at Carnegie Mellon University and high school student members of the TeleCommunity Studio at Duquesne University, Pittsburgh.

## Credits:

Robert Dunn: Project Design and Direction  
Patrick McKiernan: 3D CG and Web Design  
Lena Zehavi: Archaeologist, IMJ Youth Wing  
Melanie Carr: Project Administration

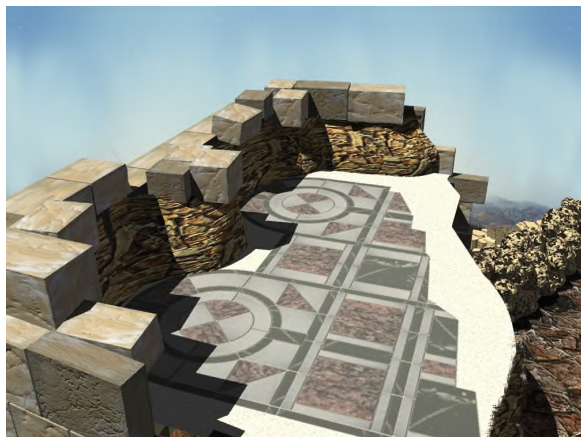




Herodian room remains with mosaic floor and columns.



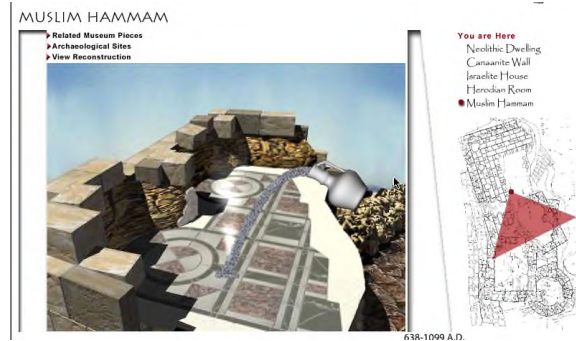
Hypothetical reconstruction of Herodian room.



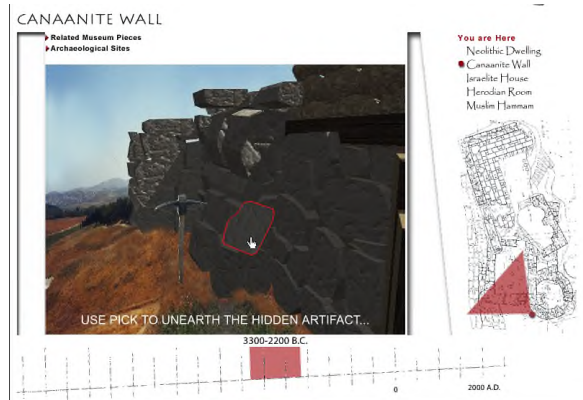
Muslim hammam, showing partial floor and wall.



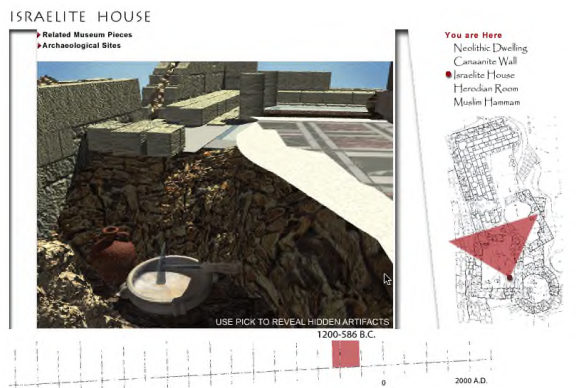
Neolithic room fire activity.



Making steam at the Muslim hammam.



Canaanite Wall, discovery of Egyptian cartouche.



Discovery of jars in the Israelite house.

# Virtual Studio: Virtual Reality in Art

DaShawn L. Hall, MA \*  
University of the Arts  
Educational Media/Media Arts

Kenneth Sakatani, Ph.D \*\*  
University of the Arts  
Educational Media - Director

## 1) Abstract

Virtual Reality or VR is a computer-designed environment that allows the user to interact within the digital environment as he or she would in a real world situation. VR has been successfully used in game development, medical technology and flight simulation. Virtual Studio is a partial immersive, interactive virtual learning environment that is based on the metaphor of the artist studio. Its primary use will be to teach elementary students about the techniques, styles and history of artist and art. Each part of Virtual Studio is named after major artists, who used or pioneered a particular style or technique. There are three core sections of Virtual Studio: 1) Architectural Design, 2) Sculptural Design, 3) Digital Painting and Color Theory.

## 2) Architectural Design: Bramante

Bramante, the Italian Renaissance Architect, was one of the top patroned artists in the 16th century. He was the chief designer of the Vatican and the Colonnade, in Rome, Italy. In this area of Virtual Studio, students will be able to act as Bramante, and will be able to physically design and create small architectural structures, such as Greek Temples, and small houses. Students will also be able to build small items, like chairs and tables to decorate their buildings.

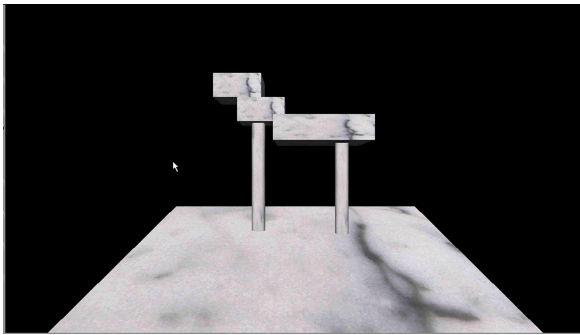


Figure 1: Image of the Bramante World

## 3) Sculptural Design: Donatello

Donatello was one of the greatest sculptors in the Early Renaissance Era. He was noted for the different mediums he used in his artwork. He also sculpted all types of objects, from figures to furniture. In this area of Virtual Studio, students will be able to experience the basic technique of sculpting. Starting from slabs of wood, stone and metal, students will be able to carve out the slabs until they form some sculptural shape, such as a figure, chair or table.

---

\* email: mstudios@bellatlantic.net  
\*\* email: ksakatani@aol.com

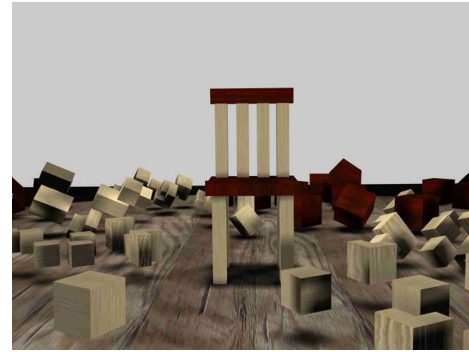


Figure 2: Image of the Donatello World

## 4) Digital Painting and Color Theory: Seurat

Seurat was the “father of Pointillism”, a technique that uses a series of dots to form a pictorial object. In this area of Virtual Studio, students will experience the basic technique of painting, by using the pointillism style. The goal is to use Particle Systems and Object Arrays, which are areas that has yet to be successful in VR world development and interactivity. Students will also learn about color theory, by developing color spectrums from primary and secondary colors.



Figure 3: Example from the Seurat World

## References

- PIMENTEL, K. AND TEIXEIRA, K. 1993. *Virtual Reality: Through the new looking glass*. McGraw Hill.
- YOUNGBLUT, C. 1998. *Educational Uses of Virtual Reality Technology*. The Institute for Defense Analyses



# An Application of Tangible Interfaces in Collaborative Learning Environments

Lori Scarlatos

Brooklyn College, Department of Computer and Information Science

2900 Bedford Ave., Brooklyn, NY 11210

lori@sci.brooklyn.cuny.edu

## 1. Introduction

Over the years, educators and government officials have searched for ways to improve learning in our schools, particularly in the areas of math and science. Many have come to recognize that collaborative activities, learning through play, and teacher guidance can help children to get over their initial fears and even begin to enjoy these subjects. Yet, at the same time, shrinking school budgets are making it harder to support these approaches to learning. Tangible Interfaces for Collaborative Learning Environments (TICLE) was conceived in response to this need [Scarlatos 2002].

TICLE embodies a different notion of support for collaborative learning. It supplements physical learning activities with computer tutors that ask relevant questions when the students get stuck. With TICLE, a group of children is given a set of physical puzzle pieces and a specific goal (such as "put these shapes together to make a square") designed to teach some math or science concept. As the children work with the puzzle, a computer system observes their actions. This system encourages the group as they make progress, and offers to give them "hints" when they don't. The hints take a scaffolding approach, asking the children to consider smaller related problems.

The advantage of TICLE is that students are allowed to focus on solving the puzzle without having to worry about how to give instructions to a machine, or whose turn it is to use the mouse. Yet they may also turn to the computer for help and further information if and when they need it. TICLE is unique in that it

- 1) makes the computer take on the role of "guide on the side" without dominating the educational activity,
- 2) allows the students to work in groups on physical learning activities, without having to learn a computer interface, and
- 3) prescribes a method for uniquely representing the state of a puzzle or model, enabling the system to rapidly check for solutions or partial solutions.

## 2. Tangram Implementation

My first prototype, which is being displayed as part of sigKIDS, demonstrates these ideas with the Tangram puzzle. The Tangram, an old Chinese geometry puzzle, consists of five triangles, one square, and one parallelogram, all precisely cut from a large square (fig. 1). One may choose to reconstruct literally hundreds of different shapes with the Tangram pieces.

The Tangram appears frequently in both elementary school math lessons and standardized tests. Very young children use it to gain familiarity with shapes, and learn how to put them together to create new shapes. Older children use it to develop a basic understanding of what "area" and "congruence" are without having to resort to formulas. Overall, the Tangram helps children to develop a geometric intuition that should help them to better grasp more complex geometric concepts later in their school careers.

## 3. TICLE Setup

My TICLE prototype consists of several pieces. Children play with the wooden puzzle on a 2'x2' Plexiglas tabletop, which they can easily gather around. A videocam, mounted below the playing surface, "watches" the puzzle pieces in play. A bank of lights around the videocam illuminates the puzzle pieces from below. Nearby, a computer monitor shows what the computer "sees". The display also provides two large buttons which may be accessed via a touch screen. The first button plays a short animation describing the object, and the rules, of the game. The second button is used to ask for a hint. A tent-like structure over the entire setup reduces interference from other light sources.

### 3.1. Tracking Puzzle Pieces

I have marked the puzzle pieces with patterns of reflective colored spots, similar to the approach taken by Underkoffler and Ishii [1998]. These patterns uniquely identify the pieces, while also indicating their locations and orientations. By viewing the pieces from below, I've virtually eliminated the problem of obscuration.

Approximately once every second, the system takes a snapshot of the puzzle's current state. After determining the position and orientation of each piece, the software finds pairs of pieces that are touching one another. Using a shorthand notation that I developed expressing how two puzzle pieces meet, the software generates a substring for each touching pair of pieces [Scarlatos 2002]. This notation has the advantage of being translation and rotation invariant, so that children working with the puzzle need not be concerned about the position or orientation of the finished puzzle. The sub-strings are sorted and concatenated in a string to produce a unique representation of the state of the puzzle. This state string may be compared to a similar string representing the solution, to check the children's progress. This state string is also examined when the children ask for a hint, to help determine what hint should be given.

### 3.2. The Interface

When children first approach the TICLE system, they are offered a menu of shapes that they can try to recreate with the Tangram pieces. They make a selection by touching the screen. This selection determines what solution is being sought and what hints are to be provided along the way. A brief animation explains what the children are expected to do.

As the children play with the puzzle, the system continuously compares the solution and current state strings. If these strings match, then the children have found the solution and congratulations are offered. If there is a partial match (found by searching for the sub-strings), then the children are making progress. Periodic encouragement is offered as more matching sub-strings are found. Finally, when the children ask for a hint (by touching the screen), the software looks for the presence (or absence) of specific sub-strings in the current state. This, along with a record of what hints have already been shown, helps to determine which hint should be given at that time.

The hints themselves are prepared as short interactive animations. A female voice (supplemented with text) asks a question about a related (but simpler) problem, while the graphics show the pieces in question. The animation then pauses, giving the children time to work out a solution. They may then touch the screen to see what the solution to the sub-problem is. Figure 2 shows a snapshot from one of the hints.

#### 4. Results

We have observed numerous groups of children playing with this prototype at the Goudreau Museum of Mathematics in Art and Science, located on Long Island, NY (fig. 3). For testing purposes, we divided large groups into pairs of children (groups of two), and assigned each pair two tasks. The first task was to make a square out of the Tangram pieces; the second task was to make another simple geometric shape (either a house or a rectangle) out of the same pieces. For the first task, half of the groups used TICLE; the control groups used a conventional Tangram with no outside help. All groups did the second task with a conventional Tangram and no help. We videotaped the children's activities and evaluated them later using Artzt and Armour-Thomas' cognitive-metacognitive framework [1992].

Not surprisingly, we found that most groups using TICLE were able to solve the first problem, while most groups using the conventional Tangram could not. Furthermore, most of the groups that solved the first problem were also able to solve the second. We also found that, in general, children using TICLE spent less time fooling around, and more time discussing approaches to the problem (with discussions frequently triggered by the hints).

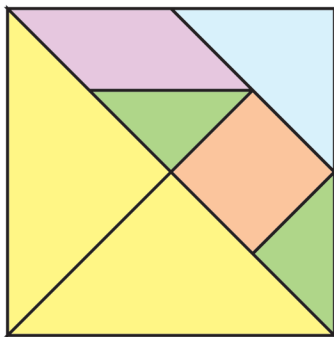


Figure 1. The Tangram, an old Chinese puzzle, can be used to make hundreds of shapes.

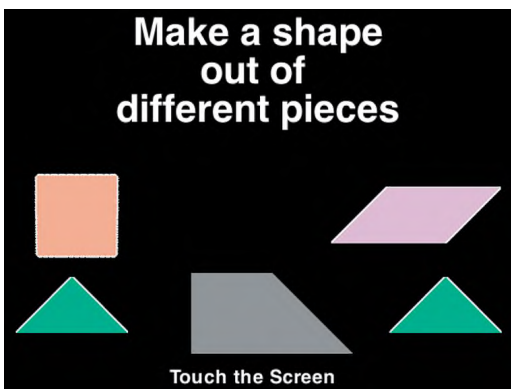


Figure 2. Screen shot of a hint given by TICLE.

#### 5. Continuing Work

I am currently working with two graduate students to develop a version of TICLE that works with a 3D puzzle. I am also working on a system that will enable teachers to define their own puzzles with hints.

#### 6. Acknowledgements

Many people helped to make this project possible: Tony Scarlatos, Shalva Landy, Saira Qureshi, Michelle Douglas, Magdalena Jaworska, Julia Breban, Robin Horowitz, Chani Sandberg, and Beth Deaner at the Goudreau Museum. This work is supported by the National Science Foundation under Grant No. 9984385, and by a PSC-CUNY grant.

#### 7. References

- ARTZT, A.F. AND ARMOUR-THOMAS, E. 1992. Development of a Cognitive-Metacognitive Framework for Protocol Analysis of Mathematical Problem Solving in Small Groups, *Cognition and Instruction* 9(2), 137-175.
- SCARLATOS, L.L. 2002. TICLE: Using Multimedia Multimodal Guidance to Enhance Learning, *Information Sciences* 140, 85-103.
- UNDERKOFFLER, J. AND ISHII, H. 1998. Illuminating Light: an Optical Design Tool with a Luminous-Tangible Interface, in *Proceedings of ACM SIGCHI 1998*, ACM Press / ACM SIGCHI, New York, 542-549.



Figure 3. TICLE at the Goudreau Museum of Mathematics in Art and Science.

# GollyGee Blocks™: A 3D Modeler for Children

Jonathan T. Blocksom\*  
GollyGee Software, Inc.†

## 1 Introduction

**GollyGee Blocks™** is a 3D modeling program for children. Children use it to stack, transform, color and texture 3D objects in a 3D scene which can be viewed from any angle. **GollyGee Blocks™** was designed as an educational, open-ended creativity tool for 3D graphics.

## 2 Program Description

Figure 1 shows a typical scene after a few blocks have been added. Children build by placing objects in the large, grass-textured square of ground in the center of the scene. The buttons at the top of the screen control menus presented on the left, which give the user different building, coloring and manipulation functions. The buttons in the lower right control the camera, which can view the scene from any angle as well as zoom in and out.



Figure 1: Sample illustration.

The child begins building by selecting 3D objects from either a menu of basic solid shapes or more complex pre-built objects. Basic solids start out with small, medium, or large sizes or sized by the user using a dragged-profile process. The Objects menu is for more complicated pre-defined objects arranged by theme; for example, people and animals, vehicles, architectural elements or castle pieces.

While placing Shapes and Objects in the scene, the user can alter the size and orientation using the keyboard's arrow keys. A toggle button at the bottom of the screen controls whether objects are initially placed directly on the ground plane or stacked up on each other. Requiring that objects are initially on the ground or stacked on each other solves the problem of multiple 3D positions being indicated by a single 2D mouse coordinate.

Once in the scene, objects can be manipulated using a set of transform tools, including translate, rotate, and uniform and non-uniform scale. Hierarchical modeling is enabled with a toggle switch, where objects are considered children of the objects they are stacked on.

Elements of the scene may also be colored or textured. A toggle is provided to allow textures and colors to be blended together (straight RGB multiplication of the texel values). A 2D backdrop

may also be colored or textured, and several of the textures are designed specifically for backdrops.

Figures 2 and 3 are scenes made with the software by students at Erving Elementary in Erving, MA. Special thanks to Michael Lipinski for providing these.

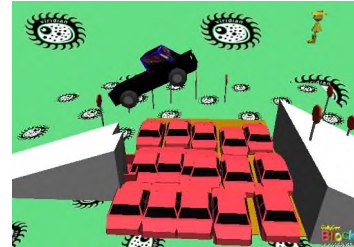


Figure 2: Car Jump by Brandon, Grade 4

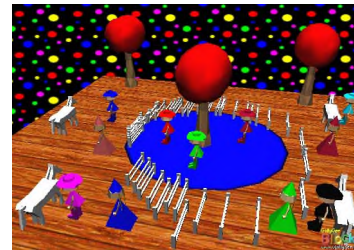


Figure 3: Rachael's Party Pool by Rachael, Grade 2

## 3 Audience and Educational Benefits

Originally written for third graders, we have found that **GollyGee Blocks™** is suitable for children in pre-school as well as sixth and seventh graders. Younger children can place blocks in the scene and experiment with coloring and texturing; older children can take advantage of the transformation tools and hierarchical modeling features to create complicated 3D scenes.

**GollyGee Blocks™** is also very effective in a classroom setting, with special appeal to visual and kinetic learners. Text and audio cues help children learn their shapes while building. More advanced concepts such as mapping skills, edges and vertices, and mental manipulation of 3D shapes are easily demonstrated with the software. Building scenes and then writing about them is also an excellent way to strengthen children's verbal-visual connection.

## 4 SigKIDS Exhibit

At sigKIDS 2002 we will be giving children and adults the opportunity to try the software as well as exhibiting scenes and pictures collected from classrooms already using it. **GollyGee Blocks™** has modest hardware requirements and will run on most personal computers purchased in the last two to four years.

\*e-mail: blocksom@gollygee.com

†http://www.gollygee.com

# Anansi's World of Folklore

Jacqueline Nuwame

## Executive Summary

Can oral storytelling live in a digital medium? Anansi's World of Folklore is a celebration of the art of storytelling, created as a broadband site for showcasing and collecting folktales. It was created to discover a way to give traditional oral storytelling a meaningful place on the Internet.

Anansi is the spinner of tales and the owner of all stories; his character serves as a departure point to the discovery of this collection of folklore tales.

The use of voice is central to Anansi; it creates an oral storytelling experience in a multimedia context. Here the user is challenged to stop and to listen and to explore their creative expression by contributing their own folktales. More than sound, Anansi offers digitally rich animation, and gives the user the selection of choice. The user decides how they will progress through the story, but the linear act of storytelling is maintained.

Within a multicultural flavor it provides an audience with a wide range of stories from many different cultures.

## Demographic

Here is a place where both parents and children can fully participate together. Parents, wanting to find a folktale from their childhood or to add one so that it can be remembered. Children can make the discovery of a variety of different stories and at the same time be encouraged to create and share their own.

Ideally targeted towards preschool children between the age range of 3 to 8. The site also has potential to reach adults who are interested in storytelling and multicultural materials. These adults would be educators, storytellers, librarians, and parents.

## Target Audience

Ideally targeted towards children between the age range of 6 to 11. Anansi also has potential to reach adults who are interested in storytelling and multicultural materials such as educators, storytellers, librarians, and parents.

## Technical Requirements

Pentium II  
Screen Display: 1024 x 768 px  
Windows 98, Flash 4 Player

## Created by:

Producers:  
Jacqueline Nuwame: [jnuwame@cdnfilmcentre.com](mailto:jnuwame@cdnfilmcentre.com)  
Judith Newman  
Art Director: Patrick Davilia

Canadian Film Center 2001



Images from Anansi's World of Folklore – Sky God



# The ToyScouts' Immersive Jukebox

## University of Central Florida

### INTRODUCTION

*The Immersive Jukebox* offers users a choice of musical experiences exploring various influences of African American Blues music: traditional African music, spirituals, work songs, and others. Inspired by the curriculum of the International House of Blues Foundation (IHOBFF)\* Blues SchoolHouse program, it introduces students and teachers to the music, art and history of the blues and its cultural origins.

The Toy Scouts are a multi-disciplinary research group made up of students, industry professionals, and research associates of the University of Central Florida that develops creative applications of virtual reality for play. The Toy Scouts were challenged with augmenting this curriculum with an affordable, mobile, highly interactive application of the material for K-12 children.



### PURPOSE

The Jukebox's selections are intended to spark self-expression as well as appreciation of musical traditions. Users are transported into a multi-sensory ambiance inspired by that which sparked the original music and art. This includes video, folk art, dance, and audiovisual effects. All the various displays work in concert to evoke performances from the users that reflect the gestural, visual, vocal and instrumental expressions of the period and genre. This mixing of visuals, music, and movement reflects the theory that most folk art is a result of deep, unique cultural collaboration and group interaction; a theory that especially applies to African American arts. For cultural and historical accuracy while working with African American folk traditions, UCF has partnered with the Department of State Division of Historical Resources Florida Folklife Program, Orange County Regional History Center and Zora Neale Hurston Museum of Fine Art.

For young users, the playful interface allows for the intuitive discovery of subtle and complex aspects of the music, empowering them to arrange unique musical hybrids of genre, musical components and audio effects as they choose. Users are encouraged to jam with virtual musicians playing traditional folk music, as they

begin to understand and explore the relationship between dance, music, and visual art. Their sessions can then be captured and played back for other visitors.



### JUKEBOX DESCRIPTION

The Immersive Jukebox Kiosk is a 10'x10'x10' booth containing a myriad of sensors that operate in parallel. A pressure-sensitive mat detects user footstep, and incorporates a full-color projected display on the same surface (the function and design of this display varies, based on the instructional intent of the program). Floor sensors allow lower-body movements to arrange musical genre, timing and elements. Image recognition detects upper-body movements, controlling audiovisual effects.

The output is equally synergistic. Video projection surfaces surround the participating group, displaying folk art influenced visuals that change as the users interact. Three-dimensional positional audio is incorporated to provide immersion and directional sources for the performances of the virtual and real musicians.

The Kiosk is also equipped with several other digital input devices, which control visual and instrumental parameters within the program. Multiple players work collaboratively with microphones, drum pads, and other instrumental input.

### DEVELOPMENT TEAM

The Immersive Jukebox was developed at the Media Convergence Laboratory, with the support of the Institute for Simulation and Training, the CREAT Digital Media Program, and the School of Electrical Engineering and Computer Science, at the University of Central Florida.

Team members include Chris Stapleton (Principal Investigator), Linda Ellis (Project Leader), Brad Martin (Music Lead), Kirsten Kischuk (Artistic Lead), Shelley Brown (Creative Investigator), Elana Rubinfeld (Artistic Expression), Matthew Gerber (Interface Engineer), Peter Stepniewicz (Technical Coordinator), John



Culbertson (Engineer), Paulius Micikevicius (Computer Vision Specialist), Kristin Congdon (Art and Cultural Advisor), Amy Hale (Folklorist).

**Christopher Stapleton**

Director of Entertainment Research,  
Media Convergence Laboratory,  
University of Central Florida  
Institute for Simulation and Training  
3280 Progress Drive, Orlando, FL 32826 USA  
+1.407.882.1356  
+1.407.658.5059 fax

\*The IHOFB is a 501 (c) (3) non-profit public charitable organization.

# Sketches & Applications – Introduction

Welcome to the Sketches and Applications Program of SIGGRAPH 2002! This program has grown from its inception to become one of the most vital parts of the annual SIGGRAPH conference. Sketches fulfill a number of different purposes:

- For the technical researchers and developers, it provides a forum for the presentation of new, novel ideas and early or incremental work.
- For people in the entertainment industry, it is an ideal place for them to explain how they did the effects in their movies and games.
- For artists, it is the place at SIGGRAPH where they can talk about the details of their art.
- For people in different disciplines, this program gives them the opportunity to show how they use computer graphics in their field.

This year we had 263 submissions from all over the world (22 different countries to be exact!) The jury accepted 149 sketches with topics ranging from rendering (photorealistic and non-photorealistic), game development techniques, virtual reality, haptics and human computer interaction, physical simulation and much, much more.

As is tradition, the Sketches program is a great place to hear about the details of blockbuster movies. We have presentations from the makers of Star Wars, Lord of the Rings, Ice Age, and many more! And for the growing audience of game developers at SIGGRAPH, we have presentations on motion capture techniques, placing the camera in games and others.

The presentations on Art, Design and Multimedia cover interacting with art, using CG for exploring art and quite a bit more. New for this year, the Sketches program is a forum for some of the Emerging Technologies installations to talk about their pieces in more details.

I owe a huge debt of gratitude to Vicki Caulfield for running and organizing the program. Her knowledge and experience are invaluable - it's nice to have a person to go to who knows everything. The Sketches Jury this year was simply terrific. I'd like to thank all my jury members and especially, the chairs of each sub-jury: Rick Parent, Linda Lauro-Lazin and Grue. Participating in the Sketches Jury is a lot of work in a short period of time, I truly appreciate the effort! Dena Slothower gets my undying gratitude for doing such an amazing job on last year's program and always being there to help me. Simon Allardice was great to hang with for a year and Stephen Spencer, Corri Nowak, Tom Rieke, Janet McAndless and, of course, Tom Appolloni do their jobs so well it was a breeze to do mine. Finally, I'd like to thank my employer, Digital Domain, for being so supportive of my work at SIGGRAPH. They have a strong commitment to the graphics community and make it easy to participate.

Doug Roble, Chair

# Sketches & Applications – Committee

## Chair

Douglas R. Roble  
Digital Domain

## Sketches & Applications Committee

Brian Barsky  
University of California, Berkeley

Juan Buhler  
PDI

David "Grue" DeBry  
Thrown Clear Productions

Steve Derrick  
Vicarious Visions, Inc.

Mark Elendt  
Side Effects Software

Darin Grant  
Digital Domain

Chuck Hansen  
University of Utah, School of Computing

Dorothy Krause  
Emeritus/Massachusetts College of Art

Linda Lauro-Lazin  
Pratt Institute

## Sketches & Applications Committee

Katerina Mania  
University of Sussex

Ioana Martin  
IBM T.J. Watson Research Center

Jacquelyn Martino  
MIT

Rick Parent  
The Ohio State University

Frédéric Pighin  
Institute for Creative Technologies - USC

Holly Rushmeier  
IBM T.J. Watson Research Center

Cheryl Stockton  
Studio Firefly/Pratt Institute

Emru Townsend  
Freelance writer, editor, and web  
designer

# 2D/3D Hybrid Character Animation on “Spirit”

Doug Cooper, Digital Supervisor  
DreamWorks Feature Animation  
coop@anim.dreamworks.com

## Abstract

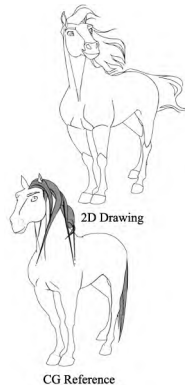
DreamWorks Feature animated film “Spirit” takes new steps in the hybridization of 2D and 3D production techniques for Character Animation, even to the extent of using both approaches within the same shot.

## 1 Introduction

The primary character in “Spirit” is a traditionally drawn 2D animated horse, which inhabits the richly detailed tapestry of the American Wild West at the turn of the century. The heavy use of 3D sets and camera moves brought an extra challenge to the traditional animators in needing to precisely match the sometimes subtle and slow perspective changes that occurred throughout a shot.

## 2 3D Reference

To help traditional animators meet this challenge, we took animated cycles of a 3D version of the character, and printed them onto animation paper. The traditional animators used these prints as perspective, volume, and scale reference when making their drawings. The animators did not rotoscope the CG renders, but instead used it solely for scale and perspective reference.



### 2.1 Digital Stunt Doubles

We modeled and setup full 3D versions of the main characters in the film to use as “digital stunt doubles” for medium to long shots, and in shots with very complicated 3D camera moves. For close-up performances, we still used the traditionally animated characters so we could accomplish the highest quality of animation performance.

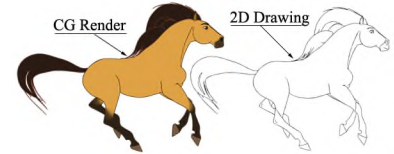
### 2.2 Take-overs

In several long and complicated shots, we used both 2D and 3D techniques within the same shot simultaneously to achieve a costs savings over animating the entire shot traditionally.

For example, in one elaborate scene, we begin with a tracking camera move following Spirit with a herd of horses. Spirit is fully 3D, as well as all the horses in the herd, and the environment. As the camera pans down and trucks into a close-up of Spirit, we “take-over” to a completely traditionally drawn version of Spirit in the middle of the shot. Spirit remains traditionally drawn for the entire close-up camera circle around him. As the camera trucks back to reveal the expansive 3D landscape, we again take-over to a fully 3D version of the character. Using this technique allowed us to achieve significant cost savings of several hundred drawings.

### 2.2.1 From 3D to 2D

We started first by animating the entire shot in CG – portions that we planned to draw were not animated fully since they were only needed for reference.



The CG animation was then rendered with a toon shader, and printed onto animation paper. The traditional animators then precisely matched several frames of the CG as drawings before taking over fully and animating the close-up. When it was time to transition back to 3D, the animators would again match the CG for a few frames. After the transition back to 3D, we kept a traditionally animated mane and tail composited onto the character for stylistic reasons.

### 2.2.2 Matching CG lines to 2D lines

After the drawings were completed and scanned in, we adjusted the toon shader parameters for the CG to more precisely match the line thickness of the scanned drawings. A simple cross-dissolve was used for the transition. We chose the transition time from 3D to 2D carefully to correspond with the character coming to an abrupt stop on a hill in the shot, helping to disguise the transition. The transition back to 3D was harder, because the movement was much slower.

### 2.2.3 Matching drawn lighting to CG lighting

On a 2D animated film, the lighting that appears on 2D drawn characters is accomplished by drawn “tone mattes” which define the areas of light and shadow. Compositors use the mattes to darken or lighten the character’s colors as desired. Traditionally drawn tone mattes are highly subjective in their design, as opposed to accurately reflecting the 3D volume of the character.

Matching these designs with 3D lighting proved to be very difficult. Whenever possible, we rendered the 3D lighting first, and asked the traditional Effects animators to match their designs to that for the transition frames. When the match wasn’t accurate enough, we used roto-mattes generated in compositing to fix the 3D renders, or the 2D drawings to bring them closer together.

## 3 Conclusions and Future Work

“Spirit” further blurs the line between animation techniques towards a future where qualifiers such as “digital” or “traditional” will be invalid, and we’ll simply have to call it “animation”. On our upcoming film “Sinbad”, we have expanded the use of Digital Characters, 3D sets and Effects, and moved to an entirely 3D layout process. We’re also developing more digital 2D tools for drawing shadows and tones directly in the computer, and training all of our traditional animators in 3D animation tools so they can apply their performance talents in the appropriate medium when needed.

# 3D Browser for Interactive Television

Valkeus Interactive Ltd  
Finland  
<http://www.valkeus.com>

Kotro Petri  
CEO  
[petri.kotro@valkeus.com](mailto:petri.kotro@valkeus.com)

Mannerkoski Olli  
VP, Research & Development  
[olli.mannerkoski@valkeus.com](mailto:olli.mannerkoski@valkeus.com)

Lesonen Hannu  
Designer  
[hannu.lesonen@valkeus.com](mailto:hannu.lesonen@valkeus.com)

Lustila Risto  
Designer  
[risto.lustila@valkeus.com](mailto:risto.lustila@valkeus.com)

## 1. Introduction

Digital media convergence is creating a vast new parallel universe, the digital dimension, where the range of available services is growing exponentially. Different forms of media and communication are integrating into consumer devices. This means new challenges for information user interfaces, especially from the point of view of user experience. The sketch introduced here is a functional prototype, a proof-of-concept, of a new kind of graphical user interface for interactive television. It integrates different services for television, Internet, radio, personal communication and games under one user interface.

## 2. 3D user interface

Interactive three-dimensional computer graphics are giving us a great potential for enhancing the user experience. Human beings are hardwired for remembering the positions of objects in three-dimensional space. Interactive 3D graphics allow us to ramp up user satisfaction by improving ease of navigation and the explorability of large information systems. Using 3D graphics in user interfaces instead of standard GUI and 2D menu systems gives us a potential to achieve more usable interaction and visual impressiveness in user interface applications. The great user experience needs both.



Screenshot from the main menu.

## 3. The structure and interaction of the prototype

The user interface is based on preview panes which are organised into the 3D space. The 3D space is a simple cylinder which user can view from inside and roll horizontally. Each preview pane is a window to “digital dimension” and depending on the subject matter holds constantly upgrading view of a television channel, game application, web-site, chat-forum or radio channel. The user can look several live channels in one view and browse through the

content that is available at the moment on the networks. By clicking the preview pane the user can enlarge the channel for full screen view and enjoy the show. The same goes of course for games, websites and other interactive entertainment.

The preview panes are divided based on the subject matter. The prototype includes categories such as pay-TV, sport, fun, people (chat), news, TV, radio, Internet, learning and recorded items. More and different categories can be easily added to the application. Inside each category can be several preview panes with live content from the networks.



Browsing the TV channels. Each preview pane shows live broadcast.



Browsing the “Fun” category. The category may include different media organised by the theme. On top of the preview panes is a banner for advertisement and below are the shortcut icons.

The prototype, 3D browser for interactive television, integrates different media forms into one simple and easy-to-use user interface where the user has always the ability to see all the content that is available at the moment on different networks. By using the potential of interactive 3D graphics we can create user interfaces that are easy to navigate and explore. And this should be our goal – to improve user experience and satisfaction.



# 3D Haptic Shape Perception Using a 2D Device

Huirong Han\*  
Ochanomizu University

Juli Yamashita†  
AIST

Issei Fujishiro‡  
Ochanomizu University

## 1 Introduction

Is a 2-dimensional (2D) force feedback device capable of presenting 3-dimensional (3D) shapes? The answer is a qualified “yes.” “Force shading”, a haptic counterpart of bump mapping in computer graphics, presents a non-flat shape on a nominally flat surface by varying the force vector direction in haptic rendering [Morgenbesser and Srinivasan 1996][Robles-De-La-Torre and Hayward 2001]. To our knowledge, such phenomena have been qualitatively measured only by 3D devices, and a quantitative comparison to 2D devices has not been made. We compare thresholds of human shape perception of the plane experimentally, using 2D and 3D force feedback devices.

## 2 Experiments

**Apparatus:** Our equipment included a Tangible Mouse (TM) (2D, Fuji Xerox Co., Ltd.) [TM URL] which conveys force to the user’s finger through a 2D flat actuator mounted on a mouse (Figure 1), and a PHANToM 1.0A (3D, SensAble Technologies, Inc.). The stimulus (Figure 2) was 2 visual control lines and a cursor that follows the haptic cursor. Haptically, a 2-polygon bump  $h$  high and  $w$  wide was located between the 2 control lines. In each trial of the experimental procedure, a pair of stimuli ( $h = 0$  and  $h > 0$ ) was presented, from which the subject was asked to choose a bump ( $h > 0$ ). Subjects were 8 adults. Four heights at regular intervals were tested, 10 times each under a different experimental condition.

**Parameters:**  $w/2$ : 10, 20 and 30 mm; surface *stiffness*: 0.4 N/mm; haptic rendering conditions: (1) normal or (2) force shading. Basically, force  $F = \text{stiffness} \cdot L \cdot \vec{N}$  was feedbacked. Under condition (1) (PHANToM alone),  $L$  is the haptic cursor penetration depth into the stimulus and  $\vec{N}$  the normal vector of the physical surface. Under condition (2), with TM, only horizontal component  $F_x$  of  $F$  ( $L$  was set to 1mm) was output. With PHANToM, output force  $F = (F_x, F_y)$ , where  $F_y = s \cdot L$  ( $L$ : penetration depth into the virtual plane) and  $F_x = F_y \cdot (2h/w)$ .

## 3 Results and Discussion

The plane shape perception threshold was estimated as a height yielding 75% correct detection for each of the subjects and conditions (Figure 3). Analysis of variance on the magnitude of “force gap” felt on the top of the bump (Figure 4) showed that (a) the difference between the two devices and widths in force shading was not statistically significant ( $p < 0.4$ ), and (b) condition (1) with PHANToM was significantly different from others and width was an effective factor ( $p < 0.05$ ). Result (a) strongly suggests that humans detect the force gap in force shading and the threshold is the same for the two devices, regardless of width or whether their coordinates are absolute (PHANToM) or relative (TM). Result (b) should represent the effect of positional feedback. Although this

research is in an early stage, the results are promising in revealing physical parameters and their thresholds that define the fidelity of haptic rendering enabling different haptic displays to be compared. Projected work includes experiments with different shapes and stiffness.

## Acknowledgements

We thank Kazushi Ozawa and the members of DPC RDC3-1G, Fuji Xerox Co., Ltd. for their support and advice. This research was funded in part by the NEDO Industrial Technology Fellowship Program.

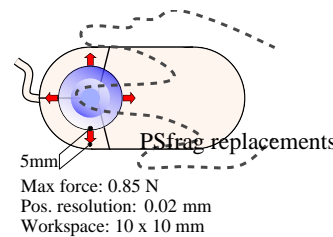


Figure 1: Tangible Mouse

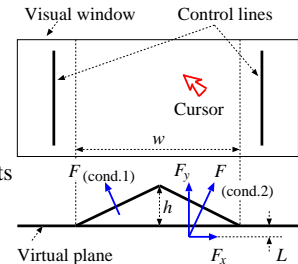


Figure 2: Stimulus

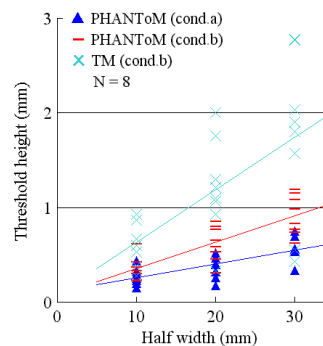


Figure 3: Results

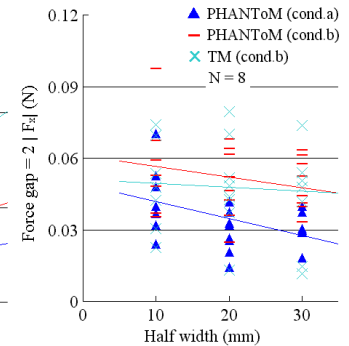


Figure 4: Max. force difference

## References

- MORGENBESSER, H., AND SRINIVASAN, M. 1996. Force Shading for Haptic Shape Perception. In *Proc. ASME Dynamics Systems and Control Division*, 407–412.
- ROBLES-DE-LA-TORRE, G., AND HAYWARD, V. 2001. Force Can Overcome Object Geometry in the Perception of Shape through Active Touch. *Nature* 412, 26, 445–448.

TM URL.  
[http://www.fujixerox.co.jp/tangible\\_mouse/](http://www.fujixerox.co.jp/tangible_mouse/)

\*han@imv.is.ocha.ac.jp  
†yamashita-juli@aist.go.jp  
‡fuji@is.ocha.ac.jp

# 3D Layout and Propagation of Environmental Phenomena for *Ice Age*

Kevin Thomason  
3D Layout Artist  
Blue Sky Studios

Robert V. Cavaleri  
Sr. Technical Director  
Blue Sky Studios

This sketch presents how the 3D Layout Department at Blue Sky Studios dealt with the production pipeline issues of transforming their first feature film *Ice Age* from sequences of hand-drawn storyboards into blocked out, ready-to-animate scenes in 3D for the Animation Department. To achieve this, custom tools were developed to automate sequence set-up and distribution of shots, in order to minimize the artists' concern with data management and maximize the creative time available to accommodate revisions in the story.

## 3D Layout

Using Maya, 3D Layout Artists were the first to visualize the storytelling of the film in 3D by blocking out the modeling, character positions, and camera animation for each sequence of shots. A shelf of MEL tools was developed to expedite the tasks of sequence set-up -- creating cameras with background storyboard panels as reference, uploading shot specific sound files, keyframing special low-resolution rigs -- such that Layout Artists could begin laying out even the most complex sequences within minutes. As the director made revisions to the timing and composition of a sequence, changes made by Layout could be updated with the touch of a button to an HTML database of sequence notes accessible to all departments of production. Additional scripts enabled Layout Artists to concatenate all the animation of the individual finalized cameras into one camera, so that other departments such as Materials and Lighting could easily hop through a sequence to see what was required for each shot.

The primary tool for shot management and distribution was the Breakout Tool, which enabled 3D Layout Artists to move quickly from shot to shot, assign shot specific geometry to a particular camera, then "break out" individual shots from the master sequence file to animators with only the essentials necessary to begin final animation. Each resulting shot file would be a streamlined version that contained only the priority geometry that characters might interact with, as well as that shot's finalized camera and blocked out characters.

## Environment Propagation

There were three basic methods that were developed to propagate vegetation over a landscape. The methods were often used in combination with one another as a 'multi-tiered' approach to achieve the desired visual complexity. Breaking the larger problem of propagation into smaller, more focused solutions also made it easier to manage.

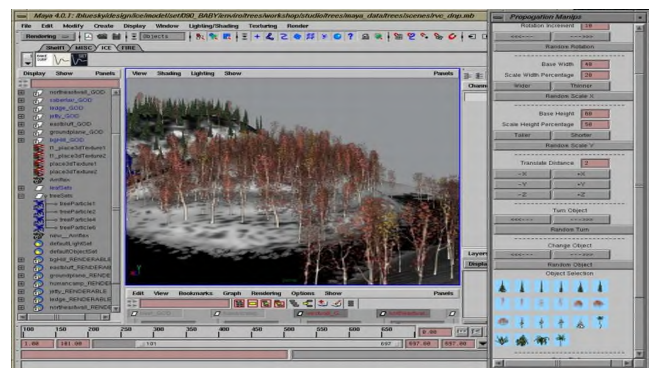
**Particle-Based Placement** - Used in cases where explicit control was needed but pre-visualization could be coarse. This method used a GUI in Maya called *propTools* and an online database containing pre-baked animation of trees & leaves. Particle data was interpreted at render time based on camera distance to

determine whether sprites or actual geometry were needed in the frame.

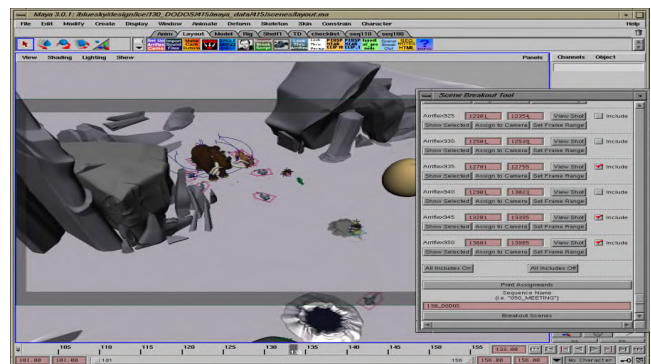
**Map-Based Placement** – Implicit placement based on texture maps. Visualization could only occur in the renderer. Used for mass placement of grass, low-level vegetation, and small pebbles.

**Geometry-Based Placement** - Manual placement of objects where maximum control was desired. Used to add enhance environmental detail in areas of the set to achieve a more organic and interesting look.

The combination of methods and tools used for environmental propagation proved to be successful in allowing us to produce acute visual detail in the environments only in the areas where it was most warranted, while filling in the rest of the environment with less expensive detail at a greater distance.



Screen Shot of the Propagation Tool



Screen Shot of the Breakout Tool

# 3D Reconstruction Of Walking Behaviour Using A Single Camera

Arunachalam Somasundaram

Rick Parent \*

## 1 Introduction

A consumer-grade camera is used to view the motion of individuals moving in an area, such as that observed in the lobby of a building. By assuming individuals are engaged in fairly normal walking behavior, we can extract sufficient information to synthetically reconstruct the scene using simple image processing techniques. The reconstructed scene can then be viewed from any angle and used to track interesting individuals. Features of the person such as shirt and pants color can be extracted from the video and applied to the synthetic model thus allowing the real individual to be recognized from viewing the reconstruction. The technique proposed has the advantage of using a single camera and would find applications in gait analysis and security. This is a work in progress and we present some initial results of a single moving figure.

## 2 Camera Calibration

A single stationary (inexpensive) webcam is calibrated using techniques proposed in [Tsai 1987] paper. The corners of a grid are marked on the floor and are used to calibrate the external parameters such as translation, rotation and also the internal focal length parameter. As the 3D reconstruction from these parameters is of reasonable accuracy in the test cases, radial distortion and scaling factors of the camera are not considered.

## 3 Capturing and Processing the Images

The scene is reconstructed from images taken using the calibrated camera. The person walking is not restricted to any path and the processing of the images uses the assumption that the person is engaged in fairly normal walking behaviour. A background image of the scene is also taken. The captured images are then processed to extract foot and head positions.

First, the background image is subtracted from the recorded images and simple thresholding produces a noisy version of the moving figure. The problem is further exacerbated by similar colors of the person and background in certain regions, by reflections, and by shadows. The resulting images are run through a connected components algorithm and the top few clusters that have an area larger than a specified threshold are chosen. The position of the resulting clusters are marked by green lines in the accompanying video. These few clusters are combined into one region and its 2D X image coordinate is calculated by averaging the positions of its clusters and is used as the X coordinate for the foot and head. The average position of this single region is marked in red. The lowest Y coordinate of the clusters is considered to be the average location of the lower foot in 2D image coordinates and the highest point that of the head.

## 4 3D Reconstruction

Using the calibrated camera parameters and by assuming that the lowest foot located in the image is in contact with the ground, the figure (actually the foot-fall) is located in 3D-space by

perspectively projecting the 2D image point to the ground plane. The height of the person is calculated by assuming that the head of the individual is, in general, directly above the average location of the individual's foot; this locates the head in 3D-space and is used to calculate the individual's height. The median of the heights extracted from all the images is applied to the model. Colors of the shirt and pants are calculated by averaging the pixels of the original image that correspond to the top and bottom portions of the clusters.

The foot locations calculated in 3D are sampled and smoothed out. A B-Spline curve is fit through these points and represents the path of the walk. A 3D model of a human is animated to walk along this calculated path. The extracted colors are applied to the shirt and pants of the synthetic model. This synthetic walk can be viewed from any angle and can be used to monitor the activity of an individual.

Currently we are working on resolving occlusion when reconstructing multiple figures and also on sampling and reconstructing textured clothes.

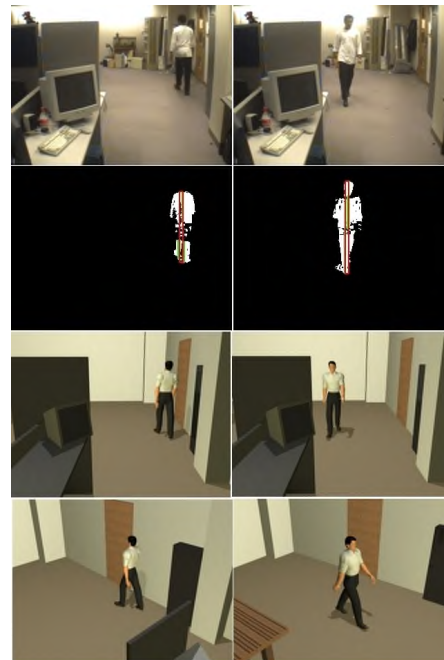


Figure 1: Actual, Processed and 3D Reconstructed Images

## References

TSAI, R. Y. 1987. A versatile camera calibration technique for high accuracy 3D machine vision metrology using off-the shelf tv cameras and lenses. *IEEE JOURNAL OF ROBOTICS AND AUTOMATON RA* - 3, 4, 323-344.

\*Dept. of Computer & Information Sciences, The Ohio State University. email: {parent,somasund}@cis.ohio-state.edu



# A 2D Sketch Interface for a 3D Model Search Engine \*

Patrick Min, Joyce Chen, Thomas Funkhouser  
Princeton University

## 1 Introduction

This sketch describes our experiences with creating query interfaces for an online search engine for 3D models. With the advent of affordable powerful 3D graphics hardware, and the improvement of model acquisition methods, an increasing number of 3D models is available on the web, creating a need for a 3D model search engine [Paquet and Rioux 2000; Suzuki 2001]. An important problem that arises for such an application is how to create an effective query interface. To investigate this issue, we created several different query interfaces, and tested them both in controlled experiments as well as in a publicly available 3D model search engine.

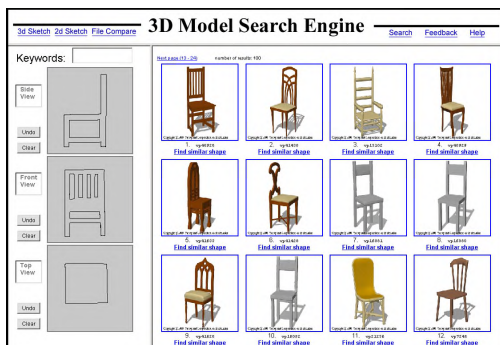
Perhaps the simplest approach to search for 3D models is to match keywords in filenames, captions, or referring webpages. However, trying to find 3D models just using text-based search techniques is not always effective: model files may be insufficiently annotated (e.g. `pr122.dxf`, `john.wrl`, misspellings, foreign languages), or ambiguously annotated (e.g. `plane.wrl`). In these cases and others, we hypothesize that shape-based queries will be helpful for finding 3D objects. For instance, shape can combine with function to define classes of objects (e.g. *round tables*).

In this sketch, we report our experiences with shape-based queries based on matching 3D models, 3D sketches, and 2D sketches.

The most straight-forward shape-based query interface is to provide the search engine with an existing 3D model and ask it to retrieve similar ones. Our search engine supports this strategy in two ways: the user may upload a 3D model file, or specify a 3D model returned in a previous search. However, a suitable 3D model may not be available to the user, in which case one has to be created from scratch. We support this by using Teddy, a simple 3D sketching tool (by T. Igarashi). Unfortunately, we found that even such a simple tool is too difficult to use for the average user. Moreover, only certain types of shapes can be created with Teddy (blobby objects with topological genus zero).

## 2 A 2D Sketch Interface

We therefore also created a 2D sketching interface that allows the user to supply up to three outline sketches as a query (see figure).



This creates the problem of how to match 2D sketches to 3D objects, which is significantly different from classical problems in computer vision: the 2D input is hand-drawn rather than photographic and the interface is interactive. We must consider several new questions: How do people draw shapes? Which viewpoints do

they select? How should the interface guide the user's input? What algorithms are robust enough to recognize human-drawn sketches? To investigate these questions, we ran a pilot study in which 32 students were asked to draw three views of 8 different objects, with a time limit of 15 seconds per object. We found that people tend to sketch objects with fragmented boundary contours and few other lines, they are not very geometrically accurate, and they use a remarkably consistent set of view directions. Interestingly, the most frequently chosen views were *not* the characteristic views predicted by perceptual psychology, but instead ones that were simpler to draw (i.e. front, side and top views).

We match the  $n$  user sketches with projected 2D images of each 3D model in the database rendered from  $m$  different viewpoints ( $m > n$ ). A model's similarity score is the minimal sum of  $n$  pairwise sketch-to-image similarity scores, subject to the constraint that no image can be matched to more than one sketch. These pairwise scores are computed by comparing their shape signatures. These signatures are based on the amplitudes of the Fourier coefficients of a set of functions obtained by intersecting the 2D Euclidean distance transform of the image with a set of concentric circles. By taking the amplitude of each coefficient, we discard phase information, thereby making the signature rotation invariant. Starting from the distance transform of the image helps make our method robust to small variations in the positions of lines.

## 3 Results

To investigate how 2D sketches can combine with text in interactive searches, we ran a user study with 43 students from an introductory computer science class. Each subject wrote up to five text keywords, and drew three simple outline sketches of 5 target objects (a chair, table, cannon, bunkbed and an elf), within a time limit of 2 minutes per object. Afterwards, we scanned their sketches and logged their keywords so that we could enter them as input to our search engine, searching in a database of 1890 household objects.

The results suggest that text and shape can be complementary in the information they provide a search engine. For example, for the chair, text keywords were not discriminating enough to differentiate it from the hundreds of other chairs and related furniture in the database: the median rank of the target chair in the text-only results was 216. Yet very simple sketches were able to describe it fairly precisely: the sketch-only median rank was 17. On the other hand, simple text keywords picked out the four bunk beds in the database (median rank 3), while the box-like sketches were not very discriminating (median rank 64). A similar effect was seen for the elf and cannon. Combining the text- and sketch scores improved the median rank to 2 for these three objects. Generally speaking, the text was effective at identifying classes of objects, while the sketches were helpful at selecting the best matches from within a class. The net result was that the combination of sketches and text usually produced a better match than either one alone.

To conclude, we believe that a 2D sketch interface is a useful query interface for a 3D model search engine. Text and shape provide different kinds of information, so used together they can be more specific.

PAQUET, E., AND RIOUX, M. 2000. Nefertiti: A tool for 3-d shape databases management. *SAE Transactions: Journal of Aerospace* 108, 387-393.  
SUZUKI, M. T. 2001. A web-based retrieval system for 3d polygonal models. *Joint 9th IFSA World Congress and 20th NAFIPS International Conference (IFSA/NAFIPS2001)* (July), 2271-2276.

\*<http://shape.cs.princeton.edu>

# A Dynamic Motion Control Middleware for Computer Games

Masaki Oshita\* and Akifumi Makinouchi  
Kyushu University, Japan

## 1 Introduction

In this sketch, we present a middleware for computer games that has the ability to realize dynamic motion control of characters. The specific novelty of the middleware is to produce dynamically changing motions in response to physical interaction between the character and environments such as collision impulses and external forces as shown in Figure 1.

In many three-dimensional computer games, it is required to produce continuous and realistic character motions on the fly. Currently most games generate such motions by combining short motion clips that are created and stored in advance. Although this approach basically works well, in this framework, characters can do nothing but just repeat prepared motion data. This limitation leaves players with a monotonous impression. The most serious problem is dynamic reactions caused by physical interaction with other characters and environments. These kinds of reactive motions should change in response to the timing, direction, and amount of the physical interaction. However, because current computer games just play a fixed motion in this kind of situation, unnatural results are sometimes generated. For example, characters fall down in the same way however it received the impact. To handle this kind of motions, the resulting motions should be generated dynamically considering the physics of the characters.

## 2 Proposed Middleware

The proposed middleware is linked and cooperates with game application. The system structure is shown in Figure 2. The application program gives an appropriate action data that is selected from the stored action data collection based on player control in context. The middleware controls the angular accelerations of the character joints based on the input action. The input action is normally executed as it is given and only when unexpected physical interaction happens, dynamically changed motions are produced. Because of this approach, the middleware makes use of existing motion clips and it is easy for the developers to use our system.

The controller in the middleware consists of three components; motion synthesizer, dynamic controller, and reaction generator. The motion synthesizer realizes smooth transition from one action to the following action considering the constraints of the end-effectors even when the character posture is slightly changed by external interaction. The dynamic controller then computes output angular accelerations for all joints based on the synthesized motion. It realizes human-like dynamic control in response to its balance and joint stresses of the character. For example, it moves the pelvis to keep its balance or swings the arms to reduce the stresses on the back or legs. Our dynamic control method is specified to human-like articulated figures and it controls each body part through a minimum number of degrees-of-freedom. The reaction generator produces dynamic reactions such as protective step for balancing, falling down and recovery motion. The reaction generator works when the dynamic controller cannot handle too large external influence. More details about the dynamic control techniques are found in [Oshita and Makinouchi 2001].

\*moshita@db.is.kyushu-u.ac.jp

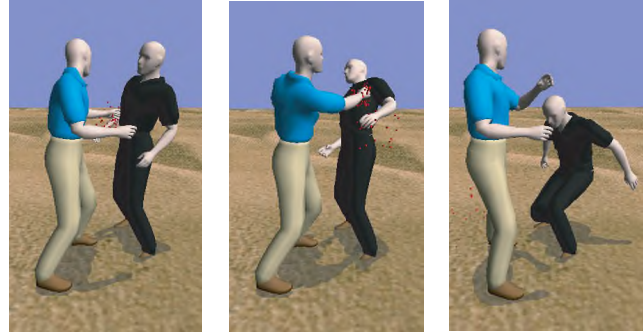


Figure 1: The middleware provides dynamic motion control in response to physical interaction between the character and environments.

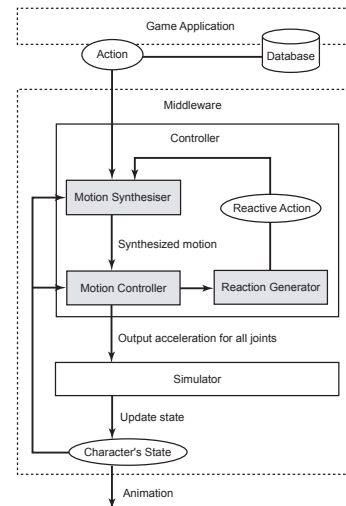


Figure 2: System structure of the middleware and application.

## 3 Application

We have developed a prototype of the proposed middleware and we are also developing a demo application which works on the prototype. The demo application is a kind of action game in which a character controlled by the user fights with other characters. Compared with existing games of similar type, it generates more appropriate dynamic motions. For example, when a character hits another, dynamic motions are generated in response to the position and power of the collisions. The developed demo application is going to be available soon on our web site: [www.e-motion.ne.jp](http://www.e-motion.ne.jp).

## References

OSHITA, M., AND MAKINOCHI, A. 2001. A dynamic motion control technique for human-like articulated figures. *Computer Graphics Forum (EUROGRAPHICS 2001)* 20, 3, 192–202.



# A Handheld Virtual Mirror

Alexandre François, Elaine Kang and Umberto Malesci  
Integrated Media Systems Center  
University of Southern California  
{alefranc,ekang,malesci}@usc.edu

This sketch presents the design and construction of a handheld virtual mirror device. The perception of the world reflected through a mirror depends on the viewer's position with respect to the mirror and the 3-D geometry of the world. In order to simulate a real mirror on a computer screen, images of the observed world, consistent with the viewer's position, must be synthesized and displayed in real-time. Our system is build around a flat LCD screen manipulated by the user, a single camera fixed on the screen, and a tracking device. The continuous input video stream and tracker data is used to synthesize, in real-time, a continuous video stream displayed on the LCD screen. The synthesized video stream is a close approximation of what the user would see on the screen surface if it were a real mirror.

## 1. Problem formulation and geometric analysis

When a person looks into a mirror, they see what they would see from the symmetrical (virtual) point of view (figure 1). A camera whose position and orientation are fixed with respect to the mirror produces a video stream that corresponds to this particular viewpoint. In order to synthesize the images corresponding to the user's viewpoint from the images captured by the camera, we must relate them geometrically. Such a transform cannot be formulated without assumptions or knowledge about the world geometry. In our case, the focus of attention is the user's own reflection in the mirror. We thus model the world as a plane parallel to the mirror, at the user's distance from the mirror. We have shown that under these assumptions, the transform relating the camera view and the user view depends only on the relative positions of the user viewpoint and the camera center, and is reduced to a scaling and a translation. The transform is exact when the user viewpoint lies on the camera axis. This model provides a straightforward and efficient way of synthesizing the viewer's image stream from the camera stream, and provides a good approximation when the viewer's position remains close to the camera axis.

## 2.2. System integration

We have implemented the mirror image transform as a module in the Modular Flow Scheduling Middleware [François 2001], the open source implementation of IMSC's software architecture for Integrated Media Systems. It is specifically designed for parallel processing of concurrent data streams in real-time, and its modularity dramatically reduces development and integration time.

The actual handheld virtual mirror system is composed of a handheld 10" LCD flat screen fitted with a camera and a magnetic tracker sensor (figure 2), a fixed processing unit, a fixed magnetic tracker transmitter, and a second magnetic tracker sensor placed on the user. Real-time camera and tracker inputs are used to synthesize the virtual mirror stream and display it in real-time on the LCD screen. The result is a very convincing handheld mirror simulation.

## 3. Perspectives

The Virtual Mirror project was initiated as an enabling technology for the Virtual Daguerreotype project for the "Lost and Found: Rediscovering Early Photographic Processes" Exhibit Museum Application [Lazzari et al. 2002]. Other mirror settings, such as a large (simulated) mirror on a wall, are currently being designed.

The design and construction of realistic virtual mirrors represents an original experiment in Human-Computer Interface, that opens the door to various interactive experiments in entertainment, arts and communication. Use of video analysis and graphics techniques will allow to explore and interfere with what has always been a private, solitary act of a person looking in a mirror.

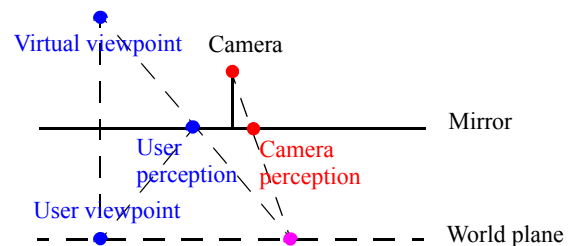


Figure 1. Basic mirror geometry.

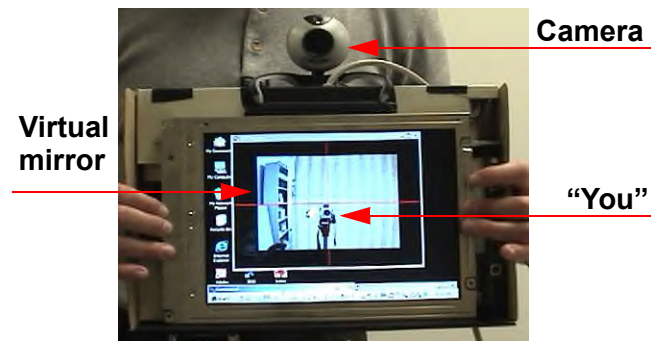


Figure 2. Handheld virtual mirror device.

## 4. Acknowledgements

Gérard Medioni, Margaret Lazzari, Margaret McLaughlin and Wee Ling Wong (USC). This research was funded in part by the Integrated Media Systems Center, a National Science Foundation Engineering Research Center, Cooperative Agreement No. EEC-9529152; and by a grant from the USC Arts Initiative.

## 5. References

- FRANÇOIS A. 2001. *Modular Flow Scheduling Middleware*. <http://mfsm.sourceforge.net/>
- LAZZARI, M., MCLAUGHLIN, M. L., JASKOWIAK, J., WONG, W., AND AKBARIAN, M. 2002. A haptic exhibition of daguerreotype cases for USC's Fisher Gallery. In *Touch in Virtual Environments: Haptics and the Design of Interactive Systems*, McLaughlin, M. L., Hespanha, J., and Sukhatme, G., Eds., IMSC Series in Multimedia, Prentice-Hall.

# A new Interface for the Virtual World Foot Motion Sensing Input Device.

Barrera Salvador, Hiroki Takahashi, Masayuki Nakajima\*  
Tokyo Institute of Technology

## 1 Introduction

Presently, most virtual reality systems use upper body parts to interact with objects in the virtual environment. This situation is caused by technological limitations of current interface devices. Starting from this viewpoint we developed a new interface for detecting ankle motions relative to the knee. We believe that hands-free navigation, unlike the majority of navigation techniques based on hand motions, has the greatest potential for maximizing the interactivity of virtual environments since navigation modes are more direct motion of the feet.

## 2 System Overview

1. Our interface consists of two rotary encoder sensors attached to a slipper for detecting the motion. The sensors are connected to a strap below the knee with rubber bands.
2. The rotary encoder sensors measure the directionality and intensity of foot motion by rotating in accordance with foot motion.

The outline of the architecture system is shown in Figure-1.

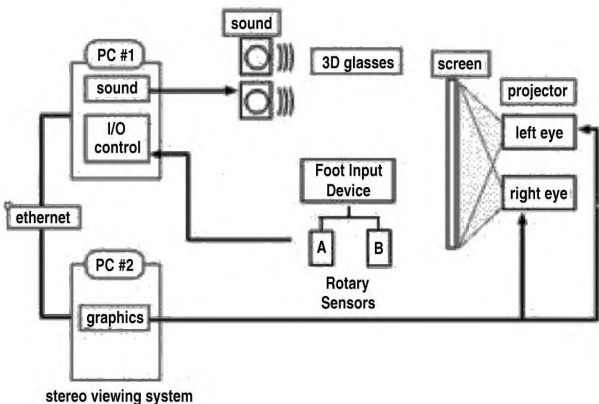


Figure 1: Overview of the architecture system.

The dimensional positions on the slipper are obtained through analog calculation electronics. The PC obtains the positions through an A/D converter.

## 3 Technology

When the users change the movement according to the motion, the sensors receive the information and directionality. Since we consider the structure of the human legs was possible to detect different movements such as figure 2.

A data transformation algorithm was made in order to calculate X and Y directional acceleration values. The basic idea of the data

transformation algorithm is a space transformation. We used the following method to define the relationship between the sensor values and the X Y values. The algorithm requires 5 key-points of basic postures (center, front, back, right and left) in the sensor space. We use stereo viewing system for displaying; correspondingly various scanning modes are being investigated, with results in number of systems incompatible to one another.

We address the problem of the interconnection of such device through standard conversions by a signal processing approach; namely the model of a universal standard converter is proposed, which is based on a layered functional architecture. The concept of a virtual standard is introduced for stereoscopic signals. In the system when this machine receives direction data, it redraws a picture according to the data such as figure 1. In summary, the system receives direction data from the slipper unit, reconstructs the scene picture and transmits it into the projector.

## 4 Results

The project presents a multi-modal interface that provides a natural space, where the user can interact with a virtual environment via the 3D graphics, audition and natural movement.

1. Replacing input with a keyboard or mouse, foot input devices systems are intuitive for current operations using hands and feet naturally in the real world.
2. The design of this interface opens the way to a new foot interface based on foot position and motion measurement.

These techniques are shown to be simple and sufficiently accurate. As well, this link technology does not disturb the visual interaction and keep enough freedom for the user.



Figure 2: Foot Input Device for the virtual world.

## References

- IWATA, H., AND FUJII, T. 1996. Virtual perambulator: A novel interface device for locomotion in virtual environment. In *Proceedings of IEEE 1996, Virtual Reality Annual International Symposium, IEEE*.
- SHINOGAWA, Y., AND YOUNG, Y. 1990. Automating view function generation for walkthrough animation using a reeb graph. In *Proceedings of Computer Animation 90*, Springer, Tokyo, 227–238.

\*e-mail: (salvador, nakajima, rocky)@img.cs.titech.ac.jp

# A Phenomenological Model for Bokeh Rendering

Juan Buhler  
jbuhler@pdi.com  
Senior FX Animator  
PDI/DreamWorks

Dan Wexler  
wex@pdi.com  
Research and Development  
PDI/DreamWorks

## Summary

This sketch describes the computer graphics techniques used to render the out-of-focus areas of a picture. We propose a simple, phenomenological model that provides control to the artist and achieves a realistic look.

## Abstract

“Bokeh” is a Japanese word used to describe the quality of the out-of-focus areas as rendered on film by a physical lens. In areas that are out of focus, the circle of confusion (i.e., the distribution of light absorbed onto film from a single point of light in the scene) is bigger than in areas that are focused. Different lenses distribute light within the circle of confusion differently, depending on a number of factors, like the shape and number of diaphragm blades and the optical design of the lens itself. For example, lenses that are designed to correct for optical aberrations tend to render a circle of confusion that is more intense on the edges and slightly less so in the center. Other lenses, like the so called “mirror” lenses, which include reflective elements besides refractive ones, typically produce donut-shaped circles of confusion since some of the light paths are cut from the center by the reflective element.

Moreover, in a real lens the circle of confusion sometimes becomes elliptical on the areas outside of the center of the image, with the shorter axis of the ellipsis oriented radially from it.

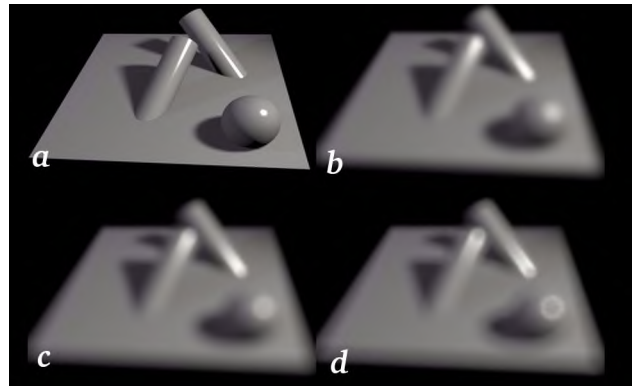
Standard rendering techniques model depth of field using point sampling techniques which converge towards a bokeh model with uniform density across the circle of confusion. This represents an idealized lens that does not exist in reality. Real lenses have non-uniform bokeh distributions. For example, a smooth, or “creamy” lens will have a Gaussian distribution across the circle of confusion. Real lenses may have different distributions at different points on the film plane.

Our technique allows the animator to specify an arbitrary probability density function to represent the distribution of intensity within the circle of confusion. The probability curve is used by the renderer to jitter the location of the sample point on the lens. For example, if the density function is a Gaussian, more samples will be taken towards the center of the lens than at the edges. Conversely, for a density function that simulates a mirror lens, fewer samples are taken near the center of the lens because the mirror blocks light as it moves through the lens. By generating enough sample points, the model converges on the true bokeh density function.

We will also show the effect of specifying two shapes for the light distribution, one for close and the other for far focused areas, and the deformation of the circle of confusion so it becomes elliptical on the edges of the image.

Future work includes the specification of the diaphragm shape, and the possible implementation of this technique as a particle renderer, which would provide efficient sampling at a much lower cost.

Additional work to add support for stratified sampling is also possible.



**Image 1:** Four versions of the same scene. a) is in focus. In b) objects are out of focus, using a Gaussian probability function. In c) the function used is a rectangular step, and in d) it is an inverted triangle. A typical modern photographic lens renders bokeh as something between c) and d)

## References

- MERKLINGER, HAROLD M. 1997 A Technical View of Bokeh. *Photo Techniques Magazine*, May/June 1997.
- ZIMMERMAN, PETER, article on bokeh. <http://www.minox.org/minoxencyclopedia/b/bokeh.html>
- WEXLER, DANIEL. *Bokeh rendering*. <http://www.flarg.com/bokeh.html>
- POTMESIL, M. AND CHAKRAVARTY, I. 1981. A lens and aperture camera model for synthetic image generation. In *Proceedings of ACM SIGGRAPH 1981*, 15 (3), pp. 297-305.
- POTMESIL, M. AND CHAKRAVARTY, I. 1982. Synthetic Image Generation with a Lens and Aperture Model. In *Transactions on Graphics*, Vol 1 No 2.

# A semiotic approach to narrative manipulation

Maria A. Alberti  
Università degli Studi di Milano  
maria.alberti@unimi.it

Dario Maggiorini  
Università degli Studi di Milano  
dario@dsi.unimi.it

Paola Trapani  
Politecnico di Milano  
paola.trapani@polimi.it

## Abstract

In the design of multimedia communication artifacts few, if any, tools support the early stage of a creative process: the heuristic project. In this work we give a proof of concept of an application addressed to a specific kind of heuristic project: given the logical sequence of episodes of a narrative, the *fabula*, the goal is to obtain different plots expressed in multi-modal language. The case study is provided by the task of transposing a written synopsis to the multi-modal language of a movie. We adopted the semiotic theory of Greimas to analyze the narrative and reveal its deep structure. The application enables users to interact with this structure in order to simulate and anticipate the effects of meaning resulting from their manipulation.

**Keywords:** Structural Semiotics, Narratology, Multimedia Communication Artifacts.

## The background

NarToo - Narrative Tool, is a prototype to be used by an author dealing with a specific kind of heuristic project, the early stage of the creative process: given the logical sequence of episodes of a narrative, the *fabula*, the goal is to obtain different *plots* expressed in multimedia language. In our case study the task is transposing a written synopsis to the multimedia language of a movie.

Semiotics discloses a very powerful perspective to look at the empirical field of design. In particular, structural semiotics applied to narratology, the tradition linking Propp to Greimas, define the elements system and the combination rules to be applied to accomplish the task of generating new plots from a *fabula*. The theory is used to analyze the narrative and reveal its deep structure. It provides a user interaction model as well, imposing the degree of freedom by defining legal and illegal operations, and suggesting an inventory of samples and hints that are helpful during the inventive process.

## NarToo and the story

The case study is provided by a synopsis based on the Borgia's, a very powerful family active in Rome. The story, starting in 1492 and extending through 1503, tells us how Rodrigo ascended the pontifical seat and managed his religious and secular power in an unscrupulous and unbiased manner. The plot, exploited several times because of its muddled nature and complex mix of human and political feelings, is related with a dark phase of the history of the papacy. The application core is fueled by the data of this story but it can be used with other subjects.

A trainee director is coping with the task of de-structuring the written text to produce new plots in a different expressive medium: for instance, a movie. The task deals with two different subtasks, addressing the different levels of *content* and *expression*. NarToo supports the early stage of the inventive process, at both levels. The trainee manipulates the chronological sequence of the events in the *fabula*, applying for instance flashback or flash-forward, in order to produce different kinds of filmic plots, assigns to actors their functional roles in each episode of the narrative and decides upon the point of view in a scene, attributing the role of

narrator to an actor. At the expression level, starting from the skeleton of the narrative substratum the author enriches the structural elements with figurative details, shifting from a mono-medium expression, which employs just the linguistic medium, to a multimedia expression, involving several perceptive registers and triggering sensorial interaction phenomena and synesthetic perceptions. Indeed, the application provides a means to interact with the narrative substratum, so that it will appear with more evidence and an archive, initially empty, where users will store their first attempts of iconographic and multimedia material to express the new plot.



The application interface provides two interaction modes: editing and browsing. A session may imply switching back and forth between the two modes. The editing mask is a matrix chart. Browsing mode enables users to sort and organize the multimedia material stored in the library during previous editing sessions, in order to analyze their poetic choices and experiment with different effects of meaning while manipulating the narrative.

The goal of film making is obviously beyond NarToo scope, that instead grants users the possibility of anticipating and simulating different effects of meaning in their initial creative process, by matching experimentally images, sounds or videos, according to their poetics. The application provides a useful tool to evaluate in advance different project solutions, preventing costly changes in later reviews.

## Implementation Notes

The core of the application is an object-oriented database, implemented in Java on Jeevan from W3apps Inc., whose objects model the elements of the semiotic analysis of the text: themes, scenes, narrative units, actors and actantial roles. The graphical interface is implemented in Macromedia Flash™.

## References

- ANCESCHI, G. Monogrammi e figure. Teorie e storie della progettazione di artefatti comunicativi. La Casa Usher, Firenze, 1981.
- GREIMAS, A. J. Du sens. Seuil, Paris, 1970.
- PROPP, V. Morphology of the Folktale. University of Texas Press, Austin, 1968.



# A Virtual Reconstruction of the Cone Sisters' Apartments

Alan Price and Dan Bailey \*  
Imaging Research Center, UMBC

## 1 Introduction

This sketch presents how the real-time 3D interactive simulation, "The Virtual Tour of the Cone Sisters' Apartments", from the SIGGRAPH 02 Art Gallery exhibition was designed, authored and produced by the Imaging Research Center (IRC) at UMBC for two installations at the Baltimore Museum of Art (BMA).



Original photograph of  
Etta Cone's Living Room  
Circa 1930s



Virtual Reconstruction of  
Etta Cone's Living Room  
(Screen grab from installation)

## 2 Background and History

Etta and Claribel Cone were two sisters who, over a period of 30 years, amassed one of the world's most acclaimed collections of early 20th century French art. This "Cone Collection", with its incomparable holdings of work by Henri Matisse and major examples of Picasso, Cezanne, van Gogh, and Renoir, was donated to The Baltimore Museum of Art along with most of the sisters' possessions and furniture in 1950. During their lives, however, the Cone Sisters lived with and displayed their collection in their apartments.

The project meticulously reconstructs the early 20th century apartments as they were and gives a glimpse of how the sisters incorporated their collection into their everyday life.

## 3 Technology and Process

The Virtual Tour of the Cone Sisters' Apartments is presented at the Baltimore Museum of Art as both a real-time interactive simulation and an immersive stereoscopic experience. In each installation, viewers navigate through a 3D model of the Cone Sisters' apartments.

Students working at the Imaging Research Center constructed the 3D model of the apartments and the furnishings inside. As is true with the development of 3D games, the models and textures had to be optimized for real-time rendering. Creative techniques were used to project shadow maps and lighting effects on the walls and floors, providing for a more photo-realistic experience not commonly seen in real-time rendering or gaming environments. Four channel spatialized audio playback tracks viewers as they move about the apartments.

Approximately 15,000 polygons make up each room. With a total of 14 rooms, the amount of data that had to be processed for real

time rendering had the potential to overload the processing capabilities of the computer. To remedy this problem, a technique was used of culling geometry not only by use of the view frustum but by tracking the camera position and culling geometry determined to be occluded by walls or other foreground objects.

The interactive navigation of the simulation was developed for a 42" flat plasma display screen with a touch sensitive overlay panel as the input device. The display is wall-mounted for permanent display in the Cone Wing of the Museum. Viewers explore and move about the apartments intuitively by touching objects, doors, and artworks.

The second version was developed to give viewers at the museum a more complete immersive experience. Driven by a network of PCs, the apartments were presented on a sixteen-foot wide by eight-foot high rear projection screen using passive stereoscopic vision and polarized glasses. Gallery visitors navigated through the apartments by using a modified joystick. The technical requirements included the need to incorporate low-cost hardware and software solutions to bring a virtual reality experience to a large public venue. The Georgia Tech Virtual Environments Group's NAVE was used as a starting point to develop our own system of a networked cluster of PCs and rear-screen passive stereo projection. The screens were constructed in a curve to partially encompass the viewers but allowing for a large standing room area. Four computers were networked together to drive the simulation; two with dual-head graphics cards to render the stereo views, one to operate as a master for positional information, audio, and input devices, and the fourth to display a graphic navigational mapping system of the apartments floorplan.



BMA gallery visitors wearing 3D glasses to view the large-scale installation. Additional monitors display a map giving the viewer's location as well as showing the original apartment photographs.

IRC Contributors: Brinton Jaecks, Sala Wong, Bea Bufrahi, Christina Hung, Christian Valiente, Ethan Berner

BMA Contributors: Allison Perkins, Jay Fisher, Katy Rothkopf

## References

PAIR ET AL. 2000. *The NAVE: Design and Implementation of a Non-Expensive Immersive Virtual Environment*. In *Conference Abstracts and Applications*, ACM Siggraph 2000, 238.

\* info@irc.umbc.edu



# Accurate Image Based Re-lighting through Optimization

Pieter Peers   Philip Dutré\*  
Department of Computer Science, K.U.Leuven

Image-based relighting represents a class of techniques that apply new lighting conditions to a scene, given a set of basis images. In this sketch we present a relighting technique that, for a single viewpoint, accurately captures the reflectance field of objects, without restrictions on their geometrical complexity or material properties. Once the reflectance field is captured, the objects can be relit under arbitrary lighting conditions. To achieve such accurate results, our method combines the strengths of both the Light Stage [Debevec et al. 2000] and environment matting [Zongker et al. 1999; Chuang et al. 2000] into a single framework.

To capture the reflectance field, the object is lit by several light patches located on a hemicycle around the object. Each patch emits a series of illumination patterns, and for each pattern, a high-dynamic range photograph is recorded. Currently, we use a calibrated CRT monitor for each hemicycle side (Figure 1).

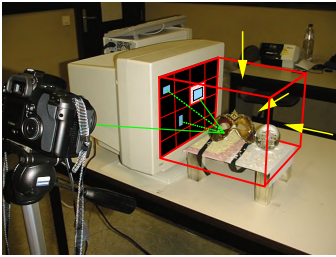


Figure 1: Current setup limits monitor placement to 4 sides, excluding the side where the camera is located. Each hemicycle side consists of a 4×4 grid of patches and we use 9 different gradient patterns per patch.

The illumination of each pixel due to emitted radiance from each light patch is approximated by a reflection coefficient and a rectangular support area on the light patch. The reflection coefficient and support are optimized through a least square minimization on the pixel radiance values.

To relight the scene using an arbitrary light map, the illumination is averaged over the support and multiplied by the reflection coefficient, for each pixel and each patch.

Due to the spacial variation of the hemicycle in light patches, diffuse materials are captured accurately. However, this subdivision is too coarse for correctly capturing specular materials. The use of illumination patterns makes it possible to find sub-patch areas which are needed for specular materials. This approach combines the advantages of both the Light Stage and environment matting, methods which handle only one of the above features adequately.

Figure 2 shows a scene, containing a decorative glass sphere and a Venetian mask, captured and relit using our framework. The results show that the proposed technique handles soft shadows correctly, specular and diffuse materials are faithfully reproduced, and even caustics (at the base of the sphere) are captured.

Our framework enables a whole new range of applications: the digital capture and display of objects with a mix of diffuse and spec-

ular appearance (e.g. jewelry or sensitive archaeological objects), augmented reality applications, (interactive) relighting of fully virtual scenes, ...

Future research will focus on faster capturing of the reflectance field, the use of other projection devices that can accommodate larger objects, and reducing the number of necessary photographs by using better illumination patterns and optimization procedures.

## References

- CHUANG, Y.-Y., ZONGKER, D. E., HINDORFF, J., CURLESS, B., SALESIN, D. H., AND SZELISKI, R. 2000. Environment matting extensions: Towards higher accuracy and real-time capture. In *Siggraph 2000, Computer Graphics Proceedings*, K. Akeley, Ed., Annual Conference Series, ACM Siggraph, 121–130.
- DEBEVEC, P., HAWKINS, T., TCHOU, C., DUIKER, H.-P., SAROKIN, W., AND SAGAR, M. 2000. Acquiring the reflectance field of a human face. In *Siggraph 2000, Computer Graphics Proceedings*, K. Akeley, Ed., Annual Conference Series, ACM Siggraph, 145–156.
- ZONGKER, D. E., WERNER, D. M., CURLESS, B., AND SALESIN, D. H. 1999. Environment matting and compositing. In *Siggraph 1999, Computer Graphics Proceedings*, A. Rockwood, Ed., Annual Conference Series, ACM Siggraph, 205–214.

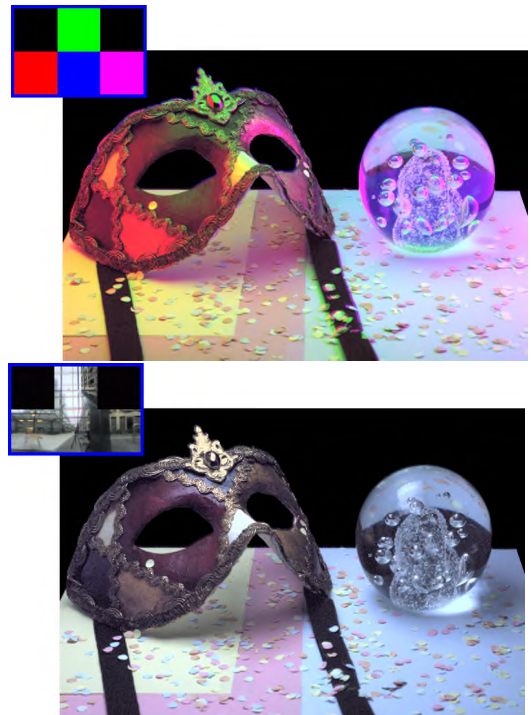


Figure 2: Venetian mask and glass sphere lit by an artificial light map and a light map taken inside a building with glass ceiling (<http://www.debevec.org/Probes>).

\*e-mail: {pieterp, phil}@cs.kuleuven.ac.be

# An Inexpensive 3D Camera

Askold Strat  
Symbol Technologies and  
SUNY at Stony Brook

Manuel M. Oliveira  
Department of Computer Science  
SUNY at Stony Brook

## Abstract

We describe the implementation of a portable 3D-camera prototype based on a consumer grade digital camera and an inexpensive laser raster generator. Such hand-held device can be used to capture smooth shapes by acquiring one or more images. Its output can be either a 3D wireframe or a textured model.

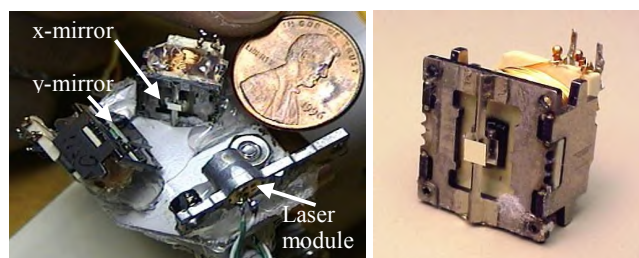
**CR Categories:** I.4.1 (Image Processing and Computer Vision): Digitization and Image Capture; B.0 (Hardware):

**Keywords:** 3D camera, 3D shape acquisition

## 1 Description

Shape acquisition has become an essential component of computer graphics systems that require modeling and visualization of objects and scenes from the real world. Most of the existing devices for capturing 3D shape have high cost and limited portability. We are building an inexpensive 3D camera using consumer grade technology to support the acquisition of smooth shapes.

Our 3D camera consists of three modules: a digital camera, a laser raster generator and some software. Although, in principle, any digital camera can be used, our current prototype uses a monochrome VGA CCD device. The laser raster generator, in turn, consists of a laser module, a high frequency oscillating mirror (x-mirror) and a low frequency oscillating mirror (y-mirror). Pictures showing the raster generator and the high frequency element are shown in Figure 1.



**Figure 1.** Raster generator (left). High frequency element (right)

The laser raster generator projects a number of parallel lines, which are deformed by the objects' surfaces (Figure 2 - left). The geometry is then recovered from the image using triangulation (Figure 2 - right). A synchronized circuit turns the laser on and off (to eliminate the horizontal retrace), depending on the movement direction of the x-mirror. A PC sound card produces the waveforms necessary to drive the two mirrors. The left and right channels of the audio output are fed to drivers of the x-mirror (2KHz) and of the y-mirror (5Hz), respectively. The faster mirror is driven with a sinusoidal waveform that matches its mechanical resonating frequency. Driving the element at resonance is necessary to maintain lower power consumption, which in our current prototype is below 0.5 W. Unlike the fast element, the y-mirror should not operate at resonance, since at resonance its movement would be sinusoidal. In order to obtain

Disclaimer: This paper reflects the personal understanding and opinions of the authors only and is not intended in any way to convey any position, policy, or opinion of Symbol Technologies, Inc. The authors and Symbol Technologies, Inc. disclaim any liability for errors or omissions in this paper.

laser planes whose angles are equally spaced, the angular velocity of the y-mirror has to be constant throughout the entire range of motion. This is not possible since the mirror and the magnet make a large mass that needs to be decelerated/accelerated when the mirror changes direction. Experimentally, we have found that the mirror's angular deflection is proportional to the current through the coil for angles up to  $\pm 17$  degrees. Thus, linear motion can be obtained away from both ends of the motion, where velocity is uniform. The camera's shutter is synchronized with the motion of the y-mirror using the mirror drive signal to trigger the shutter. The exposure time is controlled independently in the camera.

Once an image containing the raster pattern (e.g., Figure 2 - left) has been acquired, it is sent to the computer (the current prototype uses a serial interface), where it is processed. The 3D-shape reconstruction can be obtained using two different approaches. If all raster lines are visible in the image, the angular value associated with each line is immediately available and reconstruction is straightforward. In order to guarantee that all raster lines are visible, one needs to constrain the working range of the device, possibly using a back plane at the maximum acceptable range. This approach reduces the hardware requirements considerably, but the software has to make more intelligent decisions to disambiguate the lines that correspond to different projection angles. For unconstrained working ranges, two images are taken in a row from the exact same position. One of the images contain only one (or a few) raster line(s) with angular value(s) known a priori, from which the angles of all remaining lines are obtained. The use of more than one line is required if the scene contains several objects at different heights.

Figure 2 (left) shows a stripe light pattern produced by the raster generator and projected onto a person's face. The image to its right shows a wire-frame reconstruction of the visible lines from a different viewpoint.

The 3D-shape acquisition system described here is subject to the same limitations as other laser-based shape acquisition approaches. Especially challenging are surface discontinuities, which can make the lines that have been projected at different angles to appear as if they are continuous. Image segmentation is currently being considered to deal with this situation. By treating continuous areas in the image as separate entities, the entire image can be reassembled from the reconstructed entities. Once the geometry has been recovered, it will be possible to use the original texture on the imaged objects and estimate the position of the light source. We are currently working on these features.



**Figure 2.** Laser raster projected onto a person's face (left). 3D reconstruction of the lines (right).

# Art and Education using Direct Manipulation of a Sensor Array

Taly Sharon  
MIT Media Laboratory  
20 Ames St  
Cambridge, MA 02139 USA  
taly@media.mit.edu

## 1. Introduction

We describe here two installations of direct manipulation systems in the art and education domain. In these installations we use a pressure sensitive computer-projected canvas for user manipulation. As the user presses the canvas, an art piece is created, or image layers are revealed.

## 2. Characteristics of the Interaction

### 1) Manual Direct Manipulation

Interaction with the pieces is achieved using direct manipulation. Users touch and press the canvas on which the system projects, and the immediate effects occur on the same surface. Using bare hands is much more intuitive to users than manipulating a mouse. This is especially true when developing systems for children.

### 2) Group Experience

Computer interactions are usually singular, that is, are designed for a single user. Replacing the mouse and keyboard with a sensor array enables group interaction. Group interaction is useful not only in art exhibitions - but also in education.

## 3. Technology

We use a sensor array from [tekscan 2002], which measures pressure over its surface. The sensors are usually used for medical and dental applications, as well as for other research tasks [Slivovsky and Tan 2000; Kapoor et. al. 2001]. We hide two sensors under a bright tablecloth (the canvas) and project on it from the ceiling. The projector is fitted and calibrated so that the projection matches the sensors' surface. Another projector is used to support lurkers. A typical installation is presented in Figure 1.

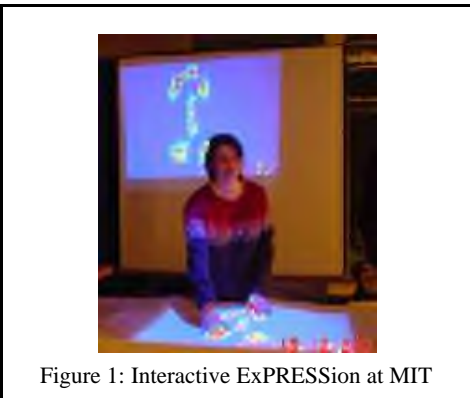


Figure 1: Interactive ExPRESSion at MIT

## 4. Interactive ExPRESSion Installation

In this installation (see Figure 1), users paint with their hands. As users press the canvas, they activate different pen colors according to the amount of pressure applied on the canvas. Several color maps as well as different interactions are supported. Examples of color maps are a blue-to-red one - that gives a feeling of warmth, and a grayscale one. Different interactions include a temporal

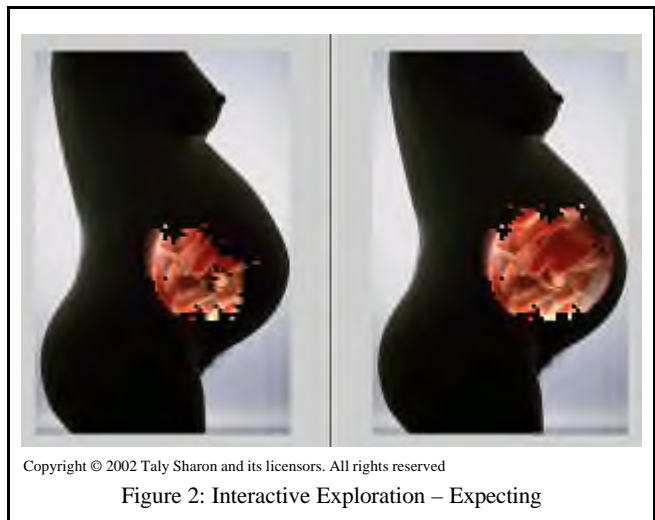
image, an accumulating image (like a real canvas) and a depreciated image in which the old details disappear over time.

## 5. Interactive Exploration Installation

In this installation, an image is first displayed on the canvas. This image appears flat; however, it is a multi-layered image, hiding other images under it. Users interacting with the canvas touch it and/or press on it. This action reveals the underlying images according to the amount of force, duration, and repetitiveness of the action. Some of the implemented scenarios are:

- ◆ *Expecting*: allows the users to view a fetus in the womb. A shadow of an expecting woman appears on the canvas. When the users press interesting areas such as the abdomen, they see the fetus hidden inside.
- ◆ *Dig Earth*: enables the users to dig into the earth and learn its layers. The dig starts from the surface, and as the diggers progress, they get deeper and deeper to view the different layers such as water, minerals, fossils, etc.
- ◆ *Go Sea*: reveals the mysteries of the sea. Starting from the blue surface, the user deepens to see the creatures underneath, until arriving at the busy seabed.

Figure 2 presents two screen shots of *Expecting*. The left one was captured at an early stage of the user interaction, while the right one comes later, as the fetus is slowly revealed.



Copyright © 2002 Taly Sharon and its licensors. All rights reserved

Figure 2: Interactive Exploration – Expecting

Similar implementations can look into a musical instrument, a car, a building, a computer program/code under its user-interface, etc.

## Acknowledgements

We thank Selene Mota for her invaluable help. Thanks also to Ted Selker for his insights.

## References

- TEKSCAN 2002, <http://www.tekscan.com/industrial/system.html>.  
SLIVOVSKY L.A. & TAN H.Z. 2000, A real-time sitting posture tracking system, Purdue Univ. Technical Report, TR-ECE 00-1.  
KAPOOR A., MOTA S. & PICARD R.W. 2001, Towards a Learning Companion that Recognizes Affect, AAAI Fall Symposium.



# ASR - Augmented Sound Reality

Daniel Dobler  
Fachhochschule Hagenberg (MTD)

Michael Haller  
Fachhochschule Hagenberg (MTD)

Philipp Stampfl  
Fachhochschule Hagenberg (MTD)

## 1 Introduction

This sketch describes the Mixed Reality application ASR (Augmented Sound Reality) which uses the overlay of virtual images on the real world to support the placement of three dimensional sound sources. Our system allows to place sound sources in a virtual or real room with the advantage of feeling, seeing and hearing them. This implies a more intuitive and better feeling of space and 3D sound.

Our main goal was to demonstrate how Augmented Reality can be used for placing 3D sound sources to enhance the impression and feeling of 3D sound. The ASR application shows how this could work in the future. Another objective of our work was to develop an AR interface (pen with marker), which allows the user to interact with the virtual objects in a very simple way. The users control the sound sources with the pen. The different sound sources are represented as 3D objects on the HMD and they can be freely positioned by the user. Moreover, the user can select a sound-type and place the sound source into the 3D space. This way several sound sources can be placed in the room and the user gets a very good impression of the behavior of sound sources in the room.

## 2 The Application

The user is able to move the sounds by moving the pen with a mounted marker in the real world. We generally decided to use a low cost setup for the application to give users the possibility to use it without special hardware. For the low cost version users only need a webcam. After starting the application, they move the markers into the web cam's field of view and place the sound sources with a simple mouse click. The high end setup includes a HMD (i-glasses), a web cam and a pen with the marker and three buttons. In this case, the webcam is mounted on the HMD and the users move the pen into their view. All the interaction is done with the pen - creating, changing, dragging & dropping and moving the sound sources into space.

The advantage of the ASR application is the direct integration and testing possibilities of 3D sound in an Augmented Reality environment. The users can directly manipulate the 3D sound source and observe the results immediately. Moreover, like in a keyframe animation it's possible to define different place points for a special sound source and in the play mode, the sound interpolates from one point to the others and finally, it's also possible to record and replay a whole session.

## 3 Implementation

The ASR project uses the ARToolKit computer vision tracking libraries [Kato et al. 1999]. The implementation of the 3D-sound is based on the Creative EAX Library [Creative 2001], which allows - besides the possibilities of OpenAL - to define various parameters for the sound sources. Moreover, the Creative EAX library supports Dolby Surround and Dolby Digital to create three-dimensional sound.



Figure 1: The ASR application.



Figure 2: Moving sound sources with ASR.

As most ARToolkit applications allow relative tracking, we implemented absolute positioning of the 3D sound objects. Using a fixed large marker in the background we easily could calculate the absolute position of the sound source. Its position is a fixed reference in the room.

## 4 Conclusion

ASR was a very interesting project combining two technologies, Augmented Reality using ARToolkit and 3D Sound Technology using the Creative EAX Library. In the future it will be more and more important to improve the realism of virtual environments integrating realistic 3D sound.

## References

- CREATIVE. 2001. EAX 2.0. Tech. rep. <http://developer.creative.com>.
- KATO, H., BILLINGHURST, M., BLANDING, B., AND MAY, R. 1999. ARToolkit. Technical report, Hiroshima City University, December.

# Automatic Generation of Pencil Drawing Using LIC

Xiaoyang Mao  
Yamanashi University  
mao@hci.media.yamanashi.ac.jp

Yoshinori Nagasaka  
Yamanashi University  
nagasaka@hci.media.yamanashi.ac.jp

Atsumi Imamiya  
Yamanashi University,  
imamiya@hci.media.yamanashi.ac.jp

## 1 Basic Idea

Line Integral Convolution (LIC)[Cabral and Leedom 1993] is a texture based vector field visualization technique. Why using LIC for pencil drawing generation? Let us look at the two images shown in Figure 1. Figure 1(a) is a digitized sample of a real pencil drawing. Look over it, we can perceive the traces of parallel pencil strokes and a gray scale tone built with the strokes. If we look at any local area of the image, however, we can find that the direction of strokes and the intensity of pixels vary randomly. The variance of intensity results from the interaction of lead material and drawing paper. The LIC image shown in Figure 1(b), however, presents the very similar features. Since an LIC image is obtained by low-pass filtering a white noise along the streamlines of a vector field, we can see traces along streamlines. On the other hand, the intensities of pixels within any local area vary randomly as the input image is a white noise. Such similarity suggests us that we can imitate the tone of pencil drawings with an LIC image.

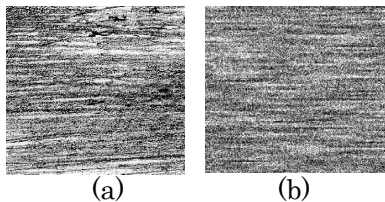


Fig.1: A pencil drawing (a) and an LIC image (b).

## 2 Technique

The largest advantage of our technique over the existing drawing generation techniques[Gooch and Gooch 2001] is that it is completely automatic. The algorithm consists of the following 7 steps:

1. Generate a white noise(Fig. 2(b)) from the input image(Fig. 2(a)). To match the tone between the input image and the resulting pencil drawing, the probability a white value is set for a pixel is proportional to the intensity level of the corresponding pixel in the input image.
2. Subdivide the input image(Fig.2(a)) into different regions(Fig. 2(c)) according to their texture or intensities.
3. Extract region boundaries (Fig. 2(d)).
4. Generate the vector field (Fig.2(e)) representing the orientation of strokes. A Fourier analysis technique is employed to extract the directions of textures. If a region does not consist of directional textures, than a randomly chosen stroke direction is assigned to it.
5. Generate pencil drawing(Fig.2(f)) by applying LIC to the white noise(Fig.2(b)) and the vector field(Fig.2(e)).
6. Add the boundary(Fig.2(d)) to the resulting image(Fig.2(f)) to obtain the drawing with outlines(Fig.2(g)).

7. Composite the resulting image (Fig.2(g)) with the paper sample (Fig.2(h)) to obtain the finished pencil drawing(Fig.2(i)).

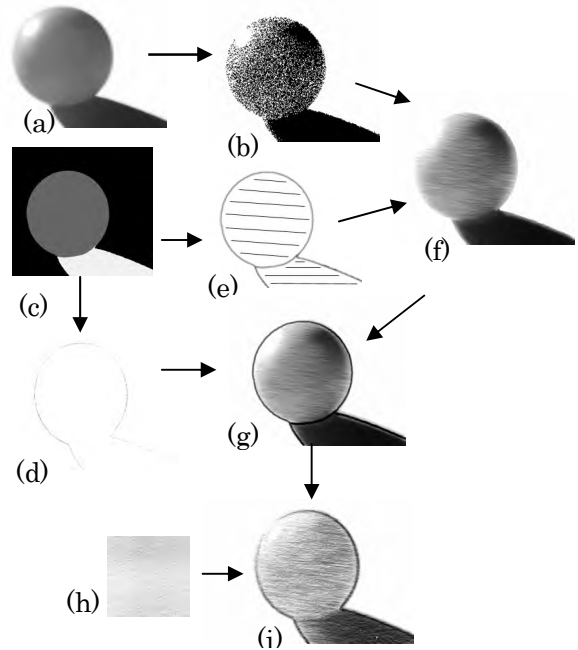


Fig. 2: Process of pencil drawing generation.



Fig.3: A pencil drawing generated with our technique.

## References

- Cabral B. and Leedom C. 1993, "Imaging Vector Field Using Line Integral Convolution", *SIGGRAPH93 conference Proceedings*, pages 263-270.
- Gooch A. and Gooch B. 2001, *Non-Photorealistic Rendering*, A.K.Peters.



## Backseat Gaming: augmented reality with speed

Liselott Brunnberg  
Mobility Studio, The Interactive Institute  
Stockholm, Sweden  
liselott.brunnberg@interactiveinstitute.se

Mark Ollila  
Norrköping Visualization and Interaction Studio  
University of Linköping  
Norrköping Campus, Sweden  
marol@itn.liu.se

This sketch presents a prototype developed as part of the Backseat gaming project. The aim of the project is to explore how to make use of mobile properties for developing compelling and fun game experiences. The prototype is developed for use in a highly mobile situation, that of a car passenger and is realized by the use of mobile devices and the users physical location during speed to merge the virtual content and surrounding road context into an augmented reality game. In this research, in addition to location, we also introduce variables such as speed, direction, timing, changing surrounding, fast movement of manipulative objects and multiple entry and exits.

The initial prototype involves a game connected to a road outside Stockholm. During the travel the player is guided by a dwarf through a dark forest inhabited with evil ghosts. The ghosts desperately seek a map that contains information about loopholes in the universe which make it possible for the ghosts to be materialized in our reality and take over the world. This map is hidden in a chip mounted somewhere in a PDA, and is the reason why the ghosts attack the player. The game content that develops while traveling along the road has clear connections to the environment seen outside the window of the vehicle, the user will for example pass a swamp, an ancient sacrifice site, and a castle that are of great importance for the game content. Except for conveying the content the mobile device can also be used as an aiming tool and radar to catch the attacking ghosts. The user can also aim it at different objects along the road, such as signs and artifacts, to pick them up and thereby use them within the game.

In this game the user is immersed in three parallel realities, the man-made environment within the car, the one outside it and the one encompassed by the virtual content. The real world passing by the vehicle acts as the world where the game takes place and is augmented with digital content, connecting the virtual and the real worlds together. These parallel realities are also connected in the opposite way where the game uses input from the real world such as artifacts and objects to enhance the game play. The narrative has to be adapted to allow different combinations of location based game events dependent on how the players enter and travels through the game. The progression of the game depends quite significantly on the speed, and hence timing is of great importance. The introduction to the game can be played out before the travel takes place, or is automatically triggered when the vehicle turns into the beginning of the road. Along the road the player will pass several places that the story of the game ties on to by triggering different location based events. These events have to be triggered in advance of the place in question and have to take into account the velocity and the direction of the car. There might for example exist various speed limits at different parts of the road or a traffic jam that slow down the speed. The speed also needs to be considered in relation to movement of manipulative objects such as the flying ghosts outside the vehicle. Other aspects that need to be examined are entry points into the game environment, that of entering the road from a point that is different then normal. The vehicle might for example enter the game from a side street, 1 kilometre into the game. Another aspect of the gameplay is, what happens if the vehicle moves back and forth, turns around and drives in the other direction or exits the game, through another side street and reenters

the game, at another location, skipping some key event that normally would have been triggered.

The Backseat gaming prototype augment the road environment with digital content to create a mixed reality combining the real surrounding with the virtual fantasy world into a compelling game. Distance has in this case been taken from a more immersed form of augmented reality using a traditional see-through head-mounted display to a more simpler form by making use of handheld mobile devices such as a context aware PDA. The game is implemented on a Compaq IPAQ Pocket PC. To receive the positioning data about the users coordinates and thereby be able to trigger different events in the game the Pocket PC is equipped with a CompactGPS card. Another important part of the game that ties the real and the virtual worlds together is for the user to be able to aim the mobile device towards a specific goal outside the vehicle. This is achieved by using a Honeywell digital compass module that provides heading, pitch and roll output mounted on the back of the Pocket PC and connected to the standard RS-232C serial port.



Figure: From top left moving clockwise. 1. Pocket PC with GPS and digital compass. 2. The castle with the ghosts overlayed. 3. The swamp with the viking and the ghosts overlayed. 4. The test environment for the game, with the targeting functioning.

For more information, visit the Mobility homepage on [www.interactiveinstitute.se/mobility](http://www.interactiveinstitute.se/mobility). Acknowledgements to Digital Bridges Ltd and the Swedish IT Institute (SIT) for funding this research. The second author performed the work whilst employed at the Interactive Institute. Thanks to Oskar Juhlin and other members of the Mobility Studio.



# Body Building Through Weight Training: Using Fitting Techniques for Skin Animation

Xiaohuan Corina Wang\* and Cary Phillips†, Industrial Light & Magic

This sketch describes a new method for deforming the skin of a digital character around its skeleton by computing statistical fit to an input *training exercise*. In this input, the skeleton and the skin move together, by arbitrary external means, through a range of motion representative of what the character is expected to achieve in practice. Using least-squares fitting techniques, we compute the coefficients, or “weights,” of our new deformation equation. The result is that the equation, which is compact and efficient to evaluate, generalizes the motion represented in the input. Once the training process is complete, even characters with high levels of geometric detail can move at interactive frames rates.

This technique is valuable because it leverages good examples of how the skin should move. These examples can come from painstaking hand sculpting by an artist, or from a muscle-based dynamic simulation process. In either case, the result is captured by an explicit deformation equation.

This technique is *pose-based*, because its input is entirely poses of the skin and skeleton, which has the advantage of being very direct. It is significantly different, however, from other techniques such as Pose-Space Deformations [Lewis et al. 2000] and Shape-by-Example [Sloan et al. 2001], because we drop the requirement that the output motion interpolate the input poses. This is beneficial because it allows a more compact representation of the motion, since the input poses are not themselves a part of the motion representation. In practice, we have found that the input poses are closely approximated.

Our new deformation equation is a generalization of a technique called *enveloping*, which defines the position of a point on the skin as a weighted combination of  $n$  limb coordinate frames:  $\mathbf{p}' = \sum_{k=1}^n w_k \mathbf{p} M_k$ . The  $M_k$  are the affine transformations of the limbs,  $\mathbf{p}$  is the rest skin point, and the  $w_k$  are the weights that determine how closely output point  $\mathbf{p}'$  follows limb  $k$ . We refer to this as Single Weight Enveloping (SWE).

Our generalized *multi-weight enveloping* (MWE) equation is an extension of this in which each entry in  $M_k$  gets its own weight:

$$\mathbf{p}' = \sum_{k=1}^n \mathbf{p} \begin{bmatrix} w_{00_k} m_{00_k} & w_{01_k} m_{01_k} & w_{02_k} m_{02_k} & 0 \\ w_{10_k} m_{10_k} & w_{11_k} m_{11_k} & w_{12_k} m_{12_k} & 0 \\ w_{20_k} m_{20_k} & w_{21_k} m_{21_k} & w_{22_k} m_{22_k} & 0 \\ w_{30_k} m_{30_k} & w_{31_k} m_{31_k} & w_{32_k} m_{32_k} & 1 \end{bmatrix}$$

The  $w_{ijk}$  are the multi-weights for limb  $k$ ,  $1 \leq k \leq n$ ,  $0 \leq i \leq 3$  and  $0 \leq j \leq 2$ . Because the original  $M_k$  is an affine transformation, its rows and columns have a geometric interpretation as direction vectors in the space of limb  $k$  with respect to the coordinate frame of  $\mathbf{p}$ . These extra weights give added controls along each dimension of each of these direction vectors. In practice, we have found this equation to be very powerful.

The input to the process consists of two parts: the poses, and *influence maps*, which identify the skin points that are affected by a particular limb, or conversely, which limbs affect a particular point. The influence maps are similar in structure to the  $w_k$  enveloping weights in the original equation, but they serve a different function here and are not a part of our resulting deformation equation.

The input poses represent samples that pair values of  $\mathbf{p}'$  with corresponding values of  $m_{ijk}$ . We solve for the weights  $w_{ijk}$  using modified least-squares fitting techniques. We also use the influence maps to scale the constraints for the solver, so that certain limb’s  $m_{ijk}$  will have more influence than others.

The new skinning process begins by preparing the input poses of the skeleton, which requires that each local region of the body be exercised through a range of motion representative of what it will achieve in practice. Then we paint the influence maps. These maps localize the effects of a pose, so one global pose can simultaneously represent poses for many sub-regions of the body.

The next step in the skinning process is to produce a proper looking skin over these skeleton, using whatever means are available. We have successfully used two techniques: hand sculpting of the model in conjunction with other simpler skin deformation techniques such as SWE, and dynamic simulation, in which the dynamics engine positions the skin automatically. In this case, we must tune the simulation to dampen the effects of inertia, so that the poses do not include shaking and jiggling.

The next step is to compute the multi-weights using the least squares solver. Once the weights have been computed, they can be validated by viewing them on the input training poses. If the results do not match, we have generally found that this points to a fault in the influence maps, which can be easily corrected.

The final step in the skinning process is to apply the multi-weights to a new sequence of animation. If it fails in certain poses, that usually indicates a deficiency in the input pose set. This failing pose can be added back to the training exercise and either corrected by hand sculpting or by rerunning another dynamic simulation, after which a new set of weights can be generated. This iterative process ensures that work done to perfect one sequence of animation successfully carries over to subsequent sequences, and the results continue to improve.

## References

- LEWIS, J. P., CORDNER, M., AND FONG, N. 2000. Pose space deformation: A unified approach to shape interpolation and skeleton-driven deformation. In *Proceedings of SIGGRAPH 2000*, ACM Press / ACM SIGGRAPH, Computer Graphics Proceedings, Annual Conference Series, ACM, 165–172.
- SLOAN, P., ROSE, C., AND COHEN, M. 2001. Shape by example. In *Proceedings of the 2001 Symposium on Interactive 3D Graphics*, ACM Press / ACM SIGGRAPH, ACM.

\*e-mail: cwang@ilm.com

†e-mail: cary@ilm.com

# Bringing Computer Graphics to Everyday Environments with Informative Art

Tobias Skog  
PLAY Studio, Interactive Institute  
tobias.skog@interactiveinstitute.se

Sara Ljungblad & Lars Erik Holmquist  
Future Applications Lab, Viktoria Institute  
{sara,l,leh}@viktoria.se

The field of display technology is rapidly developing, and LCD- and plasma-displays are already invading our surroundings. Alternative technologies such as “electronic ink”, electro-luminescent materials, and even color-changing textiles [Holmquist and Melin 2001] will further increase the number of possibilities to integrate computer graphics in our everyday lives. We believe that computer graphics for everyday life will have requirements that are very different from those of a Web page or a movie special effect. To explore this, we have developed a type of applications that anticipates a future use of computer graphics, so-called *Informative Art*.

A piece of Informative Art is a dynamic visualization – reminiscent of the style of a certain painter – whose visual appearance continuously changes, reflecting some source of information. We have previously visualized data such as e-mail traffic, current weather conditions around the world, earthquake data and the activity level in a room, in styles inspired by painters such as Piet Mondrian, Bridget Riley, Andy Warhol and Mark Rothko. Informative Art is typically shown on a wall-mounted display to give the impression of an ordinary painting. (See **Figure 1**.) Our long-term goal is that informative art will take a similar place in our daily environment as traditional artworks do today, while at the same time providing information that is relevant for the place it is situated in.

Previously, we have mostly displayed Informative Art in exhibition contexts, such as SIGGRAPH Emerging Technologies [Skog et al. 2001]. As a first step towards our goal of bringing computer graphics into everyday environments, we decided to display informative art in a public area, where we could get general perceptions and opinions from as many people as possible. We chose the Göteborg IT University, where about 150 students are present every day, as the setting for our study.

We first conducted a pre-study to establish that the piece of informative art would display data of interest for the students. Weather forecasts was one of the most common suggestions, therefore we chose to design a weather display. We decided to use Dutch painter Piet Mondrian as a source of inspiration. The simple, yet appealing, graphical structure of his compositions seemed suitable for displaying information in an intelligible way. The resulting visualization was shown at a large flat-panel screen in an open public space at the University, during one week.

The weather display looks superficially like a Mondrian painting, but in fact each colored square represents the weather of one day. (See **Figure 2**.) The size of each square reflects the temperature for that day, so that the higher the temperature, the larger the square. The color shows the weather condition: yellow represents a sunny day, blue represents rain, and red represents clouds. The display is read left-to-right, top-to-bottom, where the first square (top-left) represents the current weather conditions, the next one tomorrow forecast, and so on. The visualization is implemented as a Java application that retrieves weather information from an online weather service. As the information on the Web is updated, the application reflects the updates dynamically. The resulting image thus reflects a four-day weather forecast, while still being reminiscent of a painting in the style of Mondrian.

Did it work? Yes, provided with one brief explanation, students could interpret the information. This was confirmed in two studies conducted in a one-week testing period, which was preceded by a brief explanation of the piece to a group of students. The two studies brought in a total of 40 questionnaires, where 15 came from students who attended the briefing. The results showed that the greater part of the students considered informative art to be an enjoyable and natural part of the surroundings. Students that did not get any explanation perceived the piece solely as art, whereas the others could benefit from it as an aesthetical weather forecast.

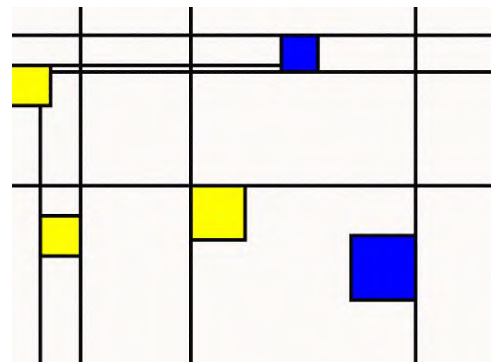
This was only a preliminary study, but the positive feedback from the test group lead us to believe that the concept of informative art is a viable way of designing new kinds of information displays, that should be investigated further. We will now design informative art for new settings, taking the next step towards integrating computer graphics in our everyday environments.

## References

- HOLMQUIST, L.E. AND MELIN, L. 2001. Using Color-Changing Textiles as a Computer Graphics Display. *Conference Abstracts and Applications of SIGGRAPH 2001*, ACM Press / ACM SIGGRAPH, New York, 272.
- SKOG, T., HOLMQUIST, L.E., REDSTRÖM, J AND HALLNÄS, L. 2001. Informative Art. *Conference Abstracts and Applications of SIGGRAPH 2001*, ACM Press / ACM SIGGRAPH, New York, 124.



**Figure 1.** Informative Art in an everyday environment



**Figure 2.** A weather forecast in the style of Mondrian



# Calming Visual Spaces: Learning from Kyoto Zen Gardens

Michael J. Lyons  
ATR Media Information Science  
mlyons@atr.co.jp

Gert J. Van Tonder  
Kyoto University  
gvtonder@yahoo.co.uk

Ian Shortreed  
ATR M.I.S.  
shortreed@mercury-soft.com

Nobuji Tetsutani  
ATR M.I.S.  
tetsutani@atr.co.jp

## 1 Introduction

Zen gardens exhibit sophisticated visual designs achieved with minimal compositions and engender a calm, contemplative atmosphere. Here we use perceptual models to study Zen garden design with the aim of discovering guidelines for the creation of calm visual spaces.

## 2 Visual Grouping

Models of scene understanding reveal how boundary contour, junction, and surface texture elements are combined by the visual system into the patterns which we perceive when looking at a scene. Gestalt laws of proximity, similarity, smoothness, and enclosedness provide an intuitive set of cues to understand this process. Analogously, in Zen gardens, visual heuristics are used to guide selection and control of visual pattern elements to express calmness and avoid conflicting, confusing scenes. Several design heuristics are at work in figure 1, e.g. the fractal-like repetition of a triangular pattern, giving this garden a harmonious, natural appearance. Further examples will be discussed in the presentation.

## 3 Medial Axis Transform

In Zen gardens considerable emphasis is placed on the beauty of “empty” space. A famous example is the dry landscape garden at Ryoanji temple (figure 2), which consists of a sparse arrangement of five clusters of 15 rocks and some moss surrounded by an expanse of raked gravel. Medial axis transform analysis, studied in models of perception and in computer vision, is useful for visualizing the structure of these empty spaces. The medial axis transform of the empty space in the Ryoanji garden reveals a non-accidental, simple, self-similar structure, which may explain the strong visual appeal of this garden.

## 4 Conclusion

Consideration of visual Gestalt cues, multi-scale patterns, and the structure of empty space deepens insight into the perceptual basis of traditional garden aesthetic qualities of asymmetry, simplicity, naturalness, and tranquility. We will describe guidelines for calming visual environments, illustrating these with photographic images and Quicktime VR [Shortreed 2001] of Kyoto’s most famous Zen gardens. These and the medial axis transform visualization technique should interest those who wish to create calm visual spaces, virtual or real.

## References

SHORTREED, I. 2001. *Kyoto Gardens: A virtual stroll through Zen landscapes*. Mercury Software CD-ROM.

**Acknowledgement** MJL and NT are supported in part by the Telecommunications Advancement Organization of Japan. GJVT is supported by the Japanese Society for the Promotion of Science.



Figure 1: Similar patterns are repeated at multiple scales in Dokuza-tei garden, Daitokuji temple.

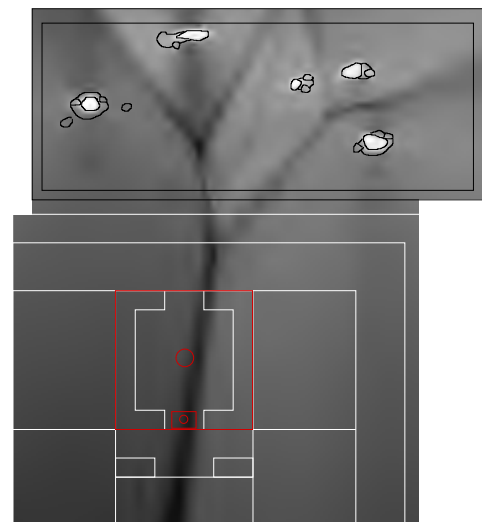


Figure 2: **Upper:** composite picture of Ryoanji garden. **Lower:** medial axis transform of Ryoanji garden with the (Edo period) temple building plan superimposed. Dark lines indicate loci of maximal local symmetry. The trunk of the medial axis, along which the view of the garden is optimal, passes close to center of the abbot’s room.

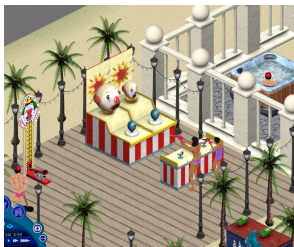


# Cameras and Point-of-View in the Gamespace

Jay Riddle\*  
Electronic Arts

## Introduction

This technical sketch surveys the wide variety of projection techniques that are employed in today's video and computer games. We will discuss the evolution of the "game view" and the reasons behind the choice of perspective, some of which are practical, while others are legacies. We will also explore the limits inherent in choosing a particular technique.



This desire for immersion has led to the use of perspective projection. This technique became feasible for application in consumer-level products in the mid 1990's, with the introduction of affordable graphics accelerator cards. Today's PC graphics cards and video game consoles such as the PlayStation2, GameCube, and Xbox have the ability to shuffle enormous volumes of data in real-time to produce stunning visuals and audio, making full first- and third-person perspective available to the masses.

Some of the questions we will explore include:

- How do we convey information to the user in a way that informs but does not interfere?
- Why are the most dramatic camera angles the worst for gameplay?
- What is a "dynamic camera" and how does it work?
- How does changing camera angles (editing) affect the game experience?
- How is editing done in a non-deterministic experience?
- Will we eventually evolve beyond the notion of a camera?

## The Evolution

When any new medium is created, it borrows from the conventions of the mediums in which it found its roots. Early video games had no notion of a camera but borrowed visual structure from the board game paradigm. All gameplay occurred in one discrete frame of two-dimensional space and resembled a plan or elevation view. Later games would begin to explore the space outside the visible window, scrolling detail by the player. These games came to be known generically as "side-scrollers".

Today's games employ several commonly used view projections: orthogonal, isometric, and perspective. All three are still widely used and one (perspective) has several flavors.

## The Choices

The choice of projection has been driven by both gameplay and technology constraints. In recent years, the technology constraint has been lessened as a third factor has crept into this decision: a desire to immerse the game player.



## The Methods

We will use imagery (both still and real-time footage) from a variety of games to underscore many of the points outlined above.

---

\*e-mail: jriddle@ea.com

# Challenges of the Homeland Pan in “Spirit”

Doug Cooper, Digital Supervisor  
DreamWorks Feature Animation  
coop@anim.dreamworks.com

## Abstract

The opening shot in “Spirit” runs a full 3 minutes without scene cuts, and introduces the audience to the breathtaking scenery of the mythic Old West. This elaborate scene was accomplished with every possible technique by mixing traditional drawing and painting art forms with 3D environments, effects, and digital characters. This Sketch will attempt to summarize some of the key challenges in producing this scene, which took over two years to complete.

## 1 Introduction

The Homeland Pan scene in “Spirit” is used as a way to welcome the audience into the film and introduce them to the vast and beautiful wild landscapes inhabited by the hero character. Originally conceived as a complex multi-plane camera move utilizing mostly 2D artwork, it quickly became evident that the true depth of the features like the Grand Canyon, Monument Valley, Yosemite, or the redwood forests couldn’t be expressed as just a series of flat panning images. Ultimately, the shot was designed with an entirely 3D layout and camera move, and we inserted 2D paintings, drawn characters, and a slew of 3D digital effects and characters into the environment to complete the effect.

## 2 Breaking it down

The animatic of the sequence was in development for approximately 9 months while the Directors and Layout leads tweaked models, compositions, and camera paths to get things just right. At a total of 4183 frames, we knew we’d need to split it into more manageable pieces for production. We went through the shot and spotted possible transitions, where we could use natural elements of the environment for wipes or to hide transitions from one set to another (such as the archway, canyon walls, forest trees, water-fall mist, and geyser spray). In some cases, we simply had to switch in mid shot, using matching portions of two separate sets to hide the transition (horses coming over the hill). Once we had broken it down into 7 distinct pieces, it was possible to begin production in the various departments.

### 2.1 3D Matte Paintings

The concept of projecting traditionally painted artwork onto 3D geometry as a form of non-photorealistic rendering is not new, but this technique has never been used in such volume in one



sequence before. The entire environment for this flythrough sequence was accomplished using a mix of 3D models with projected paintings and 2D multi-plane backgrounds. Custom tools had to be developed to streamline the process so that we could paint, project, and blend together over 1800 separate pieces of painted artwork. Approximately another 700 separate pieces of flat 2D paintings were assembled together to add trees, skies, and mountains.

Maya was used for the modeling of all geometry, creating the camera move, and setting up projectors for the 2D paintings. Paintings were executed in traditional media first, scanned into the computer, and finished in Photoshop before being projected onto the geometry and rendered in RenderMan.

### 2.2 CG Effects

To complete the detail of the natural environments, digital effects were used throughout the sequence for rushing water, grass, light rays, dust, splashes, and the occasional geyser. Traditional drawings were used to enhance the digital effects through use of our proprietary Spryticle system.



We used the CG water systems developed for the raging river rapids sequence in the film to realize the idyllic waterfall environment. Maya particles were used for the geyser effects, hoof splashes, and to animate the opening clouds. Maya was also used extensively for the dust, and for CG grass in the forest section.

### 2.3 CG & Traditional Characters

In keeping with the enhancements made to digital character use in other portions of the film, the Homeland Pan used fully digital characters for the Eagle, buffalo, and all of the running horses. Traditional drawing approaches were used for the deer, bears, and salmon. Despite the heavy use of CG characters, there were still over 2500 individual drawings created to realize the 2D characters in various parts of the sequence.

## 3 Conclusions

The Homeland Pan is a culmination of several different art forms into a single integrated whole. It represents a landmark achievement in animated filmmaking, and is the result of thousands of hours of effort by some of the most talented people working in this industry.

# Color Transformation Based on the Basic Color Categories of a Painting

Youngha Chang  
Tokyo Institute of Technology  
Graduate School of information  
Science and Engineering

Suguru Saito  
Tokyo Institute of Technology  
Precision and Intelligence Lab

Masayuki Nakajima  
Tokyo Institute of Technology  
Graduate School of information  
Science and Engineering

## 1 Introduction

Each painter renders a painting in her own style. This style can be distinguished by looking at elements such as motif, color, shape deformation and texture. Previously, [Hertzmann et al. 2001] suggested a method for applying the texture of an image to a photograph. In this paper, we will focus on the element of “color”.

Basically, color is applied depending on the subject motif. However, painters do not choose precisely the same color as that of the motif. Imagine an artist painting a forest. If she paints it in green-yellow, we may perceive the forest as being vivid. On the other hand, if she paints it in dark-olive-green, we may perceive it as being calm. In this way, by managing its saturation, brightness and hue, the artist can control viewer’s perception of her painting. Consequently, we propose a method which applies the color features of a given painting to a photograph, in order to give a similar impression. [Reinhard et al. 2001] suggested a method whose goal is similar to ours, but our method is based on color categorization characteristics of human vision, and can do all processing automatically even if the two input images have no similarity in composition.

## 2 Color Transformation Technique

While humans have an outstanding ability to discriminate between colors, it is reported that human color vision can group similar colors into the same category [Berlin and Kay 1969]. The same source notes that in many languages, there are 11 basic color categories. In this paper, we assume that if a color is converted to another color within the same basic color category, the result will not cause much incongruity in human vision. Based on this assumption, our algorithm transforms colors according to the following process:

First, the algorithm calculates a distribution of pixel color values of a photograph in  $CIE L^*a^*b^*$  color space. Then the color space is automatically segmented into 11 categories based on experimental results of the ratings of the basic color categories for each test color [Uchikawa et al. 1994]. After that, convex hull  $ch_i$  for color  $i$  ( $i = 1, \dots, 11$ ) are generated for each category. The convex hulls  $ch'_i$  for the painting are generated in a similar manner.

Next, for every pixel color value of the photograph, the algorithm finds its corresponding color in the reference painting. For a pixel color value  $p$  in  $ch_i$  of the photograph, its corresponding color value  $p'$  is given by:

$$p' = \frac{\text{Distance}(c_i, p)}{\text{Distance}(c_i, b)} (b' - c'_i) + c'_i \quad (1)$$



Figure1 : (a) target photograph

(b) reference painting

(c) result image

where  $c_i$  and  $c'_i$  are the center of gravity of  $ch_i$  and  $ch'_i$  respectively,  $b$  is the intersection point between  $ch_i$  and a line which start at  $c_i$  and goes through  $p$ ,  $b'$  is the intersection point between  $ch'_i$  and a line which start at  $c'_i$  and whose direction is from  $c_i$  to  $p$ .

By substituting each pixel color value with its corresponding color in the reference painting, we can produce an image whose color features are similar to those of the reference painting.

## 3 Result and Discussion

Figure 1 shows an example of our color transformation algorithm. Figure 1(a) shows a target photograph, 1(b) shows a reference painting, and 1(c) is the color-transformed image resulting from (a). Because our algorithm calculates a matching color point for every image pixel instead of substituting each segmented area to one focal color, the clouds and leaves preserve their color variety. Furthermore, because each pixel remains in the same color category after conversion, we realize color transformation without any sense of incongruity in spite of a large transformation of color.

In our future work, we will consider the use of white balance adjustment, the color transformation of a pixel whose color category is not present in the reference painting, and color gradation which covers two color categories.

## Acknowledgement

We would like to thank Professor Keiji Uchikawa at Tokyo Institute of Technology with providing us for the color category data.

The reference image is from *Mark Harden's Artchive*.

## References

- BERLIN, B., AND KAY, P. 1969. *Basic Color Terms: Their Universality and Evolution*. University of California Press.
- HERTZMANN, A., JACOBS, C. E., OLIVER, N., CURLESS, B., AND SALESIN, D. H. 2001. Image analogies. In *SIGGRAPH2001 Proceedings*, ACM, 327–340.
- REINHARD, E., ASHIKHMIN, M., GOOCH, B., AND SHIRLEY, P. 2001. Color transfer between images. In *Computer Graphics and Applications*, IEEE, 34–41.
- UCHIKAWA, K., KURIKI, I., AND SHINODA, H. 1994. Expression of color appearance in aperture and surface color modes with a category rating estimation method. *J. Illum. Engng. Inst. Jpn* 78, 2, 41–51.



# Computer Generated Clay Animation

Daisuke Goto  
University of Tsukuba  
dai@esys.tsukuba.ac.jp

Junichi Hoshino  
University of Tsukuba/PRESTO,JST  
jhoshino@esys.tsukuba.ac.jp

## Abstract

In this sketch, we propose a new technique for computer generated clay animation. Unlike the traditional approaches based on physical simulations, we focus on generating various animation effects produced by the clay animator.

**CR Categories:** I.3.5 (Computer Graphics): Computational Geometry and Object Modeling – Hierarchy and geometric transformation ; I.3.7 (Computer Graphics): Three-Dimensional Graphics and Realism - Animation

**Keywords:** clay animation, motion effects, deformation

## Retargetting Animation Process

Clay animation is popular in movies and games because of its hand-made flavor. However, making clay animation is very time consuming for modifying models and capturing scenes at every frame. Our approach of generating clay-like animation is to sample the creative process, and retargetting them to the approximate character model. In the current system, we focus on the following two animation effects: 1)Temporal surface deformation by fingers, 2)Generating expressive body motion.

## Generating Temporal Surface Deformation

One of the typical animation effects of clay animation is non-uniform surface deformation by pressing clay surfaces by creator's fingers. To generate temporal surface deformations, we approximate finger's shape by a mixture of gaussian function. The spacial distribution of the finger location can be determined by sampling real clay animation.

## Generating Expressive Body Motion

We classify the typical body motion effects of the clay characters into the following four categories: 1)Anticipation,2)Moving hold,3)Exagerations,4)Follow through. We sample the timing and the degree of the deformation from animation, but apparent motion is different. For example, "moving hold" effects are much longer than the conventional CG animation.

## Generating Expressive Body Motion

We classify the typical body motion effects of the clay characters into the following four categories: 1)Anticipation,2)Moving hold,3)Exagerations,4)Follow through. We sample the timing and the degree of the deformation from animation, but apparent motion is different. For example, "moving hold" effects are much longer than the conventional CG animation.



Figure3: Simulating temporal surface deformations by fingers

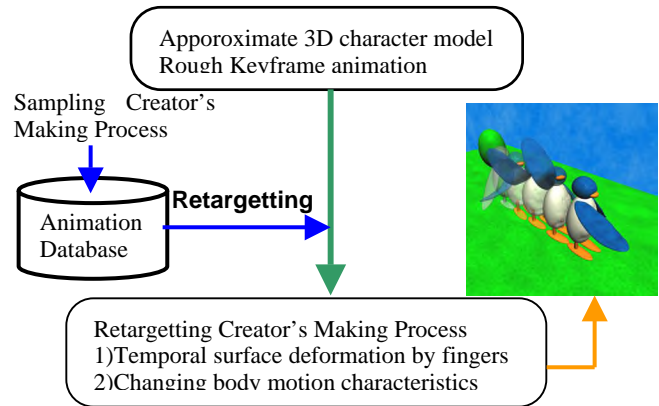


Figure1: System overview

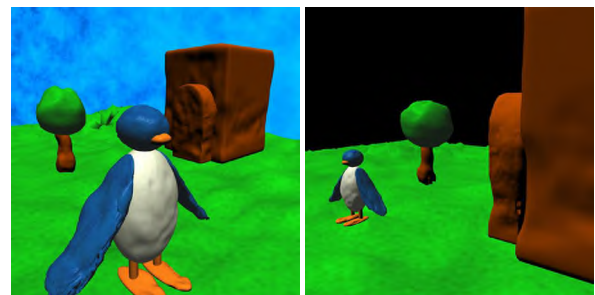


Figure2: Example of generated clay-like appearances

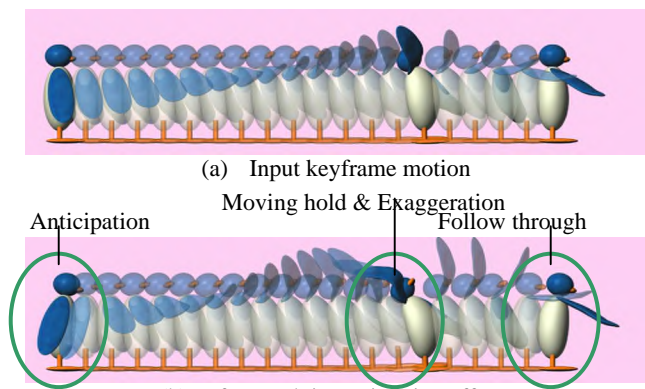


Figure4: Generating expressive motion

## References

- SHIMADA, K., MA 1993. Automated Trangulation of Surfaces and Volumes via Bubble Packing. Ph.D.thesis, Massachusetts Institute of Technology, Cambridge, MA.
- W.WELCH, AND A. WITKIN, 1992. Variational surface nokeling, Computer Graphics and Applications, vol.17, no.5, pp.40-46.

# Computer Graphics to Illustrate the Development of a Human Embryo for Professional Medical Education

Koh Kakusho<sup>1</sup> Yutaka Minekura<sup>1</sup> Michihiko Minoh<sup>1</sup> Shinobu Mizuta<sup>2</sup> Tomoko Nakatsu<sup>3</sup> Kohei Shiota<sup>3</sup>

Academic Center for Computing & Media Studies<sup>1</sup> Graduate School of Informatics<sup>2</sup> Graduate School of Medicine<sup>3</sup>  
Kyoto University, Kyoto 606-8501, Japan e-mail: kakusho@media.kyoto-u.ac.jp

## 1 Introduction

This sketch describes three dimensional (3D) computer graphics (CG) produced to illustrate the development of a human embryo for education in embryology, which is one of the basic subjects in professional medical education (Fig.1). Although similar CG have already been produced for TV programs, they are insufficient in precision for professional medical education.

In order to produce CG with sufficient precision, we make the best possible use of 3D images of real specimens taken by magnetic resonance imaging (MRI), as well as expertise in embryology. Our CG is highly valued in the community of medicine and receiving a lot of requests for distribution. From this year, we use this CG for the education in the medical school of our university.

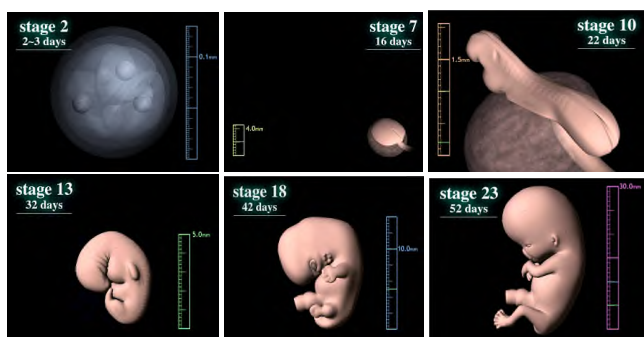


Figure 1: Illustrating the development of a human embryo.

## 2 Education in Embryology

A human embryo changes its shape drastically from an oosperm into a fetus to take the human shape while increasing the size from around 5mm to 3cm for the first eight weeks after fertilization. This process is divided into 23 periods called “Carnegie stages” (“Stages” in short)[R.O’Rahilly and F.Muller 1987]. Professional education in embryology begins with learning the standard normal 3D shape of the human embryo at each Stage.

## 3 MRI of Real Specimens

Recently, it becomes possible to obtain MRI images of human embryos[B.R.Smith et al. 1999]. Our university has more than 40,000 specimens of human embryos from Stage 13 to 23 of their development (Fig.2(a)). We selected 11 specimens from this collection so that each specimen represents a different Stage of the normal process of development, and took their MRI images<sup>1</sup>. Each image describes one of the specimens by  $128 \times 128 \times 128$  voxels with the cubic size of  $50 \times 50 \times 50$  micrometers (Fig.2(b)).

<sup>1</sup>The MRI images were taken with collaboration of Professor Katsumi Kose at the University of Tsukuba, Japan

## 4 Modeling Human Embryos

There are some difficulties in producing precise CG from MRI images. There is no specimen before Stage 12. The specimens have individual differences in shape, as well as fractures by research examinations and degradation by long preservation. MRI images have noise, distortions and limitation for the resolution. As the result, the outmost isosurface of each MRI image does not describe the complete standard shape of the human embryo (Fig.2(c)).

Since some of the authors are researchers in embryology, we carefully produced 3D models that describe the standard shapes of the human embryo with sufficient precision from the MRI images, based on our expertise in embryology. We first produced initial models by removing noise and distortions, and estimating missing features in the outmost isosurface of each MRI images (Fig.2(d)). The final 3D models were obtained by normalizing individual differences in the initial models. In the example of Fig.2, the positions of joints and the width of limbs of the model in (d) do not actually describe the standard shape. The result after normalizing these features is shown in (e). For the Stages without specimens, we produced models from scratch based on our expertise in embryology.

The process of modeling from the MRI images becomes more troublesome for the inside of the embryo. We are currently producing 3D models for the inside as our ongoing work (Fig.2(f)).

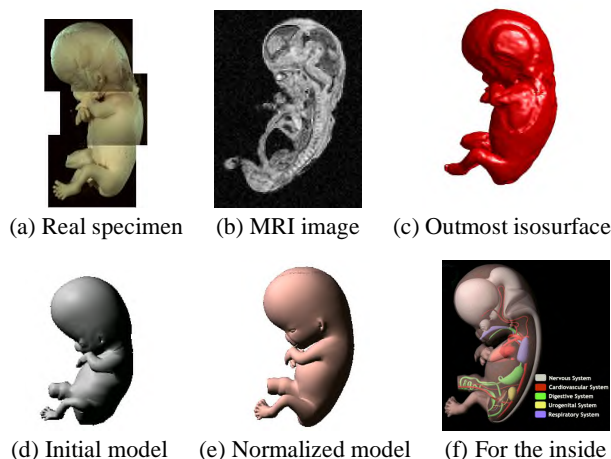


Figure 2: Modeling from a real specimen.

## References

- B.R.SMITH, D.S.HUFF, AND G.A.JOHNSON. 1999. Magnetic resonance imaging of embryos: An internet resource for the study of embryonic development. *Computerized Medical Imaging and Graphics* 23, 33–40.
- R.O’RAHILLY, AND F.MULLER. 1987. *Developmental Stages in Human Embryos*. Carnegie Institution of Washington.



# Creating 3D Painterly Environments for Disney's "Treasure Planet"

Kyle Odermatt  
Walt Disney Feature Animation  
Kyle.Odermatt@disney.com

Chris Springfield  
Walt Disney Feature Animation  
Chris.Springfield@disney.com

Since the first color short, Walt Disney animated films have been known for their beautiful and breathtaking environments that seamlessly integrate with the hand drawn characters that populate them. Over the years, technological innovations have streamlined the process of creating these environments while still maintaining the extremely high artistic standards that have made our films famous. Other advances, like the multi-plane camera, have helped to revolutionize how the camera moves through these environments. In general, such technological innovations are inspired by an artistic need which must be achieved without being prohibitively expensive. On Walt Disney's "Treasure Planet", new innovations to the DeepCanvas renderer were inspired by such artistic needs and desires.

The development of the DeepCanvas process for "Tarzan" allowed the directors to explore the jungle environment using dynamic 3D camera and prop movements that would have been impossible to achieve using traditional 2D paintings. Prior to DeepCanvas, integrating 3D environments and props into our films while maintaining continuity between the 3D rendered scenes and the surrounding 2D painted scenes was quite difficult. When it was necessary, it was achieved with limited success using texture maps and procedural shaders. For "Tarzan", the painterly art direction and the organic environment of the jungle coupled with the more extensive use of 3D camera movements called for a different approach. Inspired by these artistic needs, Eric Daniels, George Katanics and Tasso Lappas developed a stroke-based rendering process called DeepCanvas. It allowed our background artists to paint the 3D environments and props using DigPaint, our proprietary painting software, as if they were just 2D paintings instead of texture maps. Each artist was able to paint lighting, color and atmospheric nuances into the jungle environments in ways that were nearly impossible or were much more expensive to achieve using other methods. For "Tarzan", as well as "Atlantis: The Lost Empire", the original DeepCanvas process easily satisfied the artistic needs of these shows without being too expensive. However, the action-adventure story in "Treasure Planet" presented new artistic challenges to the integration of 2D and 3D environments which inspired further innovations to the DeepCanvas renderer and process.

The art direction of "Treasure Planet" follows a painterly style inspired by the "Brandywine School" of painters. This style makes strong use of light and shadow. Areas in light have much greater detail and color saturation while areas in shadow tend toward general shapes and subdued colors. In films like "Tarzan", the lighting is fairly static within each shot which enables the background artists to "paint in the lighting". Unfortunately, "Treasure Planet" did not have that same luxury. Many shots called for either dynamic lighting of an environment or elements of the environment to move in and out of the light. In the past, lighting changes like this were done using mattes that could either color correct the affected areas or reveal a separate "lit" version of the painting. However, the number of shots requiring this type of effect and the 3D nature of many of the environments made such an approach impractical.

"Treasure Planet" also had a significant number of different environments and props. Some of these environments would be used a few times while others would be re-used extensively. Prior to "Treasure Planet", most of the DeepCanvas environments and props were usually painted for each shot. Given the number of environments and props that would need to be painted and re-painted, this approach simply wouldn't be possible if we were going to use DeepCanvas.

In order to tackle these major challenges, we developed a new lighting model, a brush stroke "shader" mechanism and a post-procedural lighting painting pass inside the DeepCanvas renderer. The lighting model is built around the Renderman paradigm of light and surface shaders. This allowed us to write a set of shaders that worked in both Renderman and DeepCanvas. Therefore, our lighters could quickly light their scenes using our Renderman-based lighting tool and then pass these lights and materials on to DeepCanvas to render the scene using the painted brush strokes. A significant innovation was to incorporate each stroke's position, size, color and type into how the surfaces reacted to the lights. By doing so we were able to render painterly shadows and highlights which were more natural than ones achieved using texture maps and procedural shaders. More importantly, this allowed our painters to influence the size and shape of these aspects of the lighting.

We also developed a hierarchical "stroke shader" approach to applying brush strokes to our individual surfaces. An artist could then paint smaller pieces of the environment which could be duplicated multiple times. Then the artist could add more brush strokes to larger groupings of the surfaces to give each of the duplicate pieces a unique look. Using this method, we were able to paint large environments with a relatively small number of brush strokes while still achieving our painterly style.

To further enhance our painted environments, we developed a shader that adjusted the surface color saturation based on the amount of light incident on that surface. We also developed the ability to paint brush strokes after the lighting was applied to the surface. Thus, the artist could add "painted in" effects like reflections and global illumination as well as minimize detail in the shadow regions. These innovations allowed us to generate images that followed the "Brandywine School" style.

We intend to give an overview of these new techniques by breaking down a series of scenes from the film. Using both still images and video loops, we will illustrate how we painted, lit and rendered scenes involving 3D camera movement, dynamic lighting, and/or 3D props. We will also discuss how we used this technique to aid in the creation of our 2D backgrounds. Finally, we will demonstrate the enormous scope of the work achieved by showing a wide variety of our final backgrounds as well as several scenes in their final form. By the end of the sketch, the audience should have a clear idea of how we developed and applied these new techniques to create the three dimensional painterly environments on "Treasure Planet".

## Overview

Our team was tasked with the project of creating a full-screen, completely believable, CG stunt double for a well-known motion picture star. Successful project completion required adapting and applying many years worth of computer graphics research in a new and novel way. Most importantly, development had to be done within usual production constraints-- ultimately guaranteed results, known pipeline, quick turnaround on tests, and a final product that would rival any work in this genre ever attempted.

### “Flexible” Skin

The development of our skin rendering system involved combining various aspects of research generated over the years in a phenomenologically accurate way that would satisfy the requirements of a fast moving production.

In concordance with accepted research on the subject, we observed that what makes CG skin believable is the layering of the effects of various properties derived from analysis of how light travels and is attenuated throughout the multiple physical layers of skin. From an artistic standpoint, adding together many subtle effects re-creating the unique reflectance properties of skin arrived at an exceedingly convincing model.

To that end, we separately analyzed the diffuse and specular properties of skin and derived differences from traditional lambertian, Phong, Oren-Nayer, etc. CG counterparts. Using the short R&D time scheduled, we deviated from these traditional methods and were able to construct and perfect a skin shader relying on subsurface scattering conclusions. Artistic control through shader parameters and various texture maps was included to facilitate the artistic process.

At a glance, the critical components to the subsurface scattering skin setup are as follows:

- Photographic techniques involving polarized light enabled our team to extract reference material detailing the composition of the diffuse and specular properties of our actor. [Debevec, et al. 2000]
- The reflectance of light through the homogeneous particle medium of both the outer epidermal and inner dermal layers of skin creates a very soft, flat effect with increased attenuation at glancing incident and viewing angles. Further, skin demonstrates strong, anisotropic forward scattering. [Hanrahan and Krueger 1993]
- Phase functions for modeling directional scattering are pivotal. We formed a base around the Henyey-Greenstein principles and built out the specular term with oil influence maps and simulation of the Fresnel effects at the oil-air interface. [Henyey and Greenstein 1941]

The “flexibility” of this approach allowed us to generate renderings so convincing, that shots originally decided to be completed through more-costly, non-CG methods were added to the list of shots completed in an all CG manner.

## Color & Lighting, Rendering

We constructed a breakdown between the aspects of skin rendering that would require a ray tracing, or Monte Carlo solution, and the aspects that could be completed through a non-raytracing renderer such as PRman. It was necessary to come up with some very new and highly analytical solutions to this breakdown as our production would need to rely as heavily as possible on the already established MTOR / PRman system used at our facility, while taking advantage of newer techniques in the areas of subsurface scattering and global illumination. To that end, we implemented subsurface scattering models that rely on the fact that the layers of skin contain a statistically homogenous distribution of light interacting particles. Due to the uniformity of the model, we were able to implement this behavior in the Rendeman Shading Language through analytic techniques.

For the components of the breakdown that required a Monte Carlo approach for physically accurate results, we designed a hybrid system that could collapse into a non-raytraced solution. For instance, accurate lighting and shadows requiring a Monte Carlo based global illumination approach were distilled down to the essence of what needed to be accomplished. In this case, the interface between the lighting and shading aspects of our skin setup required input of incident hemispherical lighting and the light that would be reflected from the air-oil boundary. To obtain these inputs, we photographed fisheye images of the lighting environments that our CG actor is rendered into. A cosine weighted integral convolution image provided incident hemispherical lighting, providing an efficient global illumination pipeline required for this production.

For some time, most of the production CG world has been rendering their elements to “screen color space” or, more precisely, viewing linear results on a non-linear output. This is usually non-problematic as long as this notion is taken into account during final output to film. However on this project, we combine non-linear texture maps with linear lighting data. This color space paradox quickly became a nightmare. In order to solve this problem, it was decided that the current color and lighting pipeline would be changed such that all render input and output would be in linear color space. Render previews would be displayed through a custom screen gamma correction LUT, preserving the full linear image data through the pipeline. Compositing would be done in a linear, floating point space with screening occurring in the same, “corrected” fashion. Through these new methods, we preserve the ideality of linear color space while also preserving a final look that more accurately represents filmed imagery.

## References

- DEBEVEC, P., HAWKINS, T., TCHOU, C., DUIKER, H., SAROKIN, W. AND SAGAR, M. 2000. Acquiring the reflectance field of a human face. *Proceedings of ACM SIGGRAPH 2000*, 145-156.
- HANRAHAN, P., AND KRUEGER, W. 1993. Reflection from layered surfaces due to subsurface scattering. *Proceedings of ACM SIGGRAPH 1993*, 165-174.
- HENEY, L.G. AND GREENSTEIN, J.L. 1941. Diffuse radiation in the galaxy, *Astrophysical Journal* 93, 70-83.



# Curvature Driven Sampling of Displacement Maps

Johannes Hirche, Alexander Ehlert\*

WSI/GRIS University of Tübingen, Germany

## 1 Introduction

Displacement Mapping [Cook 1984] is commonly used in commercial rendering software for adding surface detail. In contrast to Bump Mapping not only the normal is perturbed, the surface is modified as well. A given point  $\mathbf{P}$  on a surface is displaced to a new point  $\mathbf{P}'$  in the direction of the surface normal of the point  $\mathbf{P}$ ,  $\mathbf{N}$  scaled with the value stored in the Displacement Map:

$$\mathbf{P}' = \mathbf{P} + \mathbf{D}(u_{\mathbf{p}}, v_{\mathbf{p}}) \cdot \mathbf{N} \quad (1)$$

The Displacement Map is applied to a coarse mesh that is re-tessellated accordingly. This re-tessellation can be performed equally for the whole mesh, resulting in a well structured mesh but also a large amount of unnecessary triangles, straining the graphics pipeline. [Doggett and Hirche 2000] proposed to adaptively re-tessellate the mesh using a set of tests on the perturbed normals and a summed area table of the Displacement Map, involving a number of complex arithmetic operations and a in modern graphics cards not readily available summed area table. To overcome these problems we propose to store one single additional bit per entry in the Displacement Map that is precomputed using a curvature test that can be implemented using advanced pixel shader technologies of recent graphics cards.

## 2 Algorithm

The decision whether to split an edge of a triangle is based on the curvature of the Displacement Map at that position. To avoid additional storage requirements on the graphics controller only a single bit per value of the Displacement Map is used. This bit can be stored together with the map, enlarging it only by one bit per entry, and reducing the texture lookups to only the one necessary to get the actual displacement value. When a triangle is being tested, a lookup at three sample points of every edge is performed, the two endpoints and the midpoint of the edge, and depending on the value of the decision bit a new vertex is inserted at the midpoint of every edge being split. New triangles are then inserted according to [Doggett and Hirche 2000] and re-tested until a maximum recursion level is reached.

### 2.1 Curvature Test

The main criteria for setting the decision bit is proximity to a region of sufficiently high curvature, thus making re-meshing in that area necessary. Although complex algorithms for calculating the splitting criterion are possible, our approach employs for this pre-computing step only features available in commodity graphics hardware. A Laplacian filter kernel was used to calculate the curvature of the Displacement Map. Since the filter is very sensitive to high frequency noise as it is common in data acquired from range scanning systems a Gaussian filter can be applied prior to calculating the decision bit, or if the computation is performed on the host, the two steps can be combined. In present graphics cards the convolution with a Gaussian filter kernel is not possible with a reasonably sized filter kernel. Once the curvature is calculated for the Displacement Map the values are compared to a given threshold value and

stored in the decision bitmap. Since point sampling is used while re-meshing the algorithm is subject to aliasing artifacts. To avoid this problem the decision bit is not only set for exactly the location of sufficiently high curvature, but for a surrounding region of a specified radius. Figure 1 shows the curvature of the Crater Lake Displacement Map with the resulting decision map.

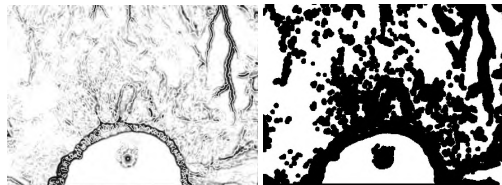


Figure 1: Curvature of the Crater Lake Displacement Map and the corresponding Decision Map.

## 3 Results, Future Work and Hardware Implications

We have implemented the described algorithm in software and the results of the re-meshing process are shown in figure 2. We are currently investigating the possibilities of storing more than one bit per sample of the Displacement Map, higher values meaning a higher curvature in the area, and re-meshing according to it and the recursion level.

Given a implementation for adaptively re-meshing as described in [Doggett and Hirche 2000] our approach would require considerably less computation in the rasterization pipeline as only one texture lookup has to be performed.

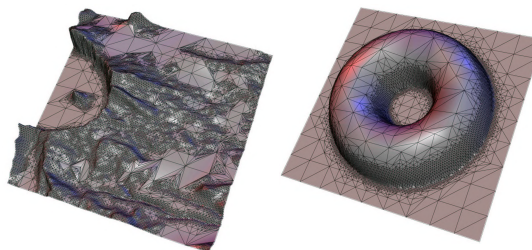


Figure 2: Result of the re-meshing process using the Crater Lake and a half-donut shaped Displacement Map.

## References

- COOK, R. L. 1984. Shade Trees. *Computer Graphics (Proceedings of SIGGRAPH 84)* 18, 3, 223–231. Held in Minneapolis, Minnesota.
- DOGGETT, M., AND HIRCHE, J. 2000. Adaptive view dependent tessellation of displacement maps. In *Proc. of Eurographics/SIGGRAPH workshop on graphics hardware 2000*, 59–66.

\*e-mail: {jhirche,ehlert}@gris.uni-tuebingen.de



## Custom Designs for Digital Imaging on Textiles

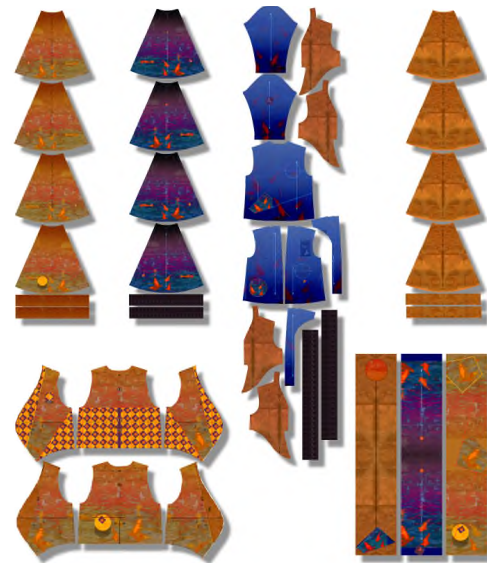
In working with custom designs for digital imaging on textiles, source images and garment patterns are scanned and the images mapped to fit the pattern design in Adobe Photoshop. For these pieces the images were scanned on the Eclipse 48" x 96" flatbed scanner.



The designs were printed onto paper-backed silk using Lyson Reactive dye in the Mutoh printer with the Wasatch SoftRip at 720 x 720 DPI.



The paper backing was removed from the silk which was then rolled in tissue paper and placed in a Jacquard steamer to heat-set the dye. After 15 minutes of steaming, the fabric was unrolled and the material washed with Jacquard Silk Wash to remove excess dye. (A stove top steamer, shown below left, is also available for heat-setting two yards at a time.) The water was squeezed out with a towel and the fabric ironed until dry. The pattern pieces were cut from the fabric and the garments were sewn. The fabrics are fully washable.



Dawn Lake Ensemble  
All Designs c2002 Lhotka All Rights Reserved



The Dawn Lake ensemble was inspired by a photograph of a duck pond taken by my son, Greg Lhotka, in Tokyo. ThunderCoat (above right) was created from one of my 48" x 96" acrylic paintings.

Today over 100 fabrics including silk, linen, wool and synthetics are being made for direct digital imaging. Reactive and acid dyes are available for both piezo and thermal printing. Recently natural fabrics have been introduced that do not require post-processing. At the present time, textile imaging is limited to fine art use, the high-end accessories market and garment sampling. As products and process are refined, digital textile imaging will be used for mass customization in many fabric applications.



# Depth-Complexity Based Occluder Selection

Gerhard Kurka

Johannes Kepler University Linz

Densely occluded regions containing many stacked objects along the line of view generally show a high local depth-complexity (DC). With respect to occlusion culling such regions require dense occluder-sets, in order to provide gap free occlusion. Using pure size or distance based selection heuristics fail to consider such regions. This sketch presents a novel dynamic occluder selection approach, which selects occluders that cover regions of high DC. Depth-Complexity Based Occluder Selection [Kurka 2001] is based on the evaluation of a low resolution (128x128 pixel) raw scene depiction (RSD), which is rendered once per frame by a simple but fast image-based rendering (IBR) technique.

During a pre-processing step, a series of low-resolution (8x8 Pixels) reference images with per-pixel depth information is created for every object in the scene. The objects are automatically created by a spatial recursive bi-section of the scene's polygons. For each object reference images are taken from 26 viewpoints. The actual RSD-rendering starts by determining the bounding rectangles of the objects' bounding boxes. Afterwards, a reference image is chosen for each object and rasterized into the corresponding bounding rectangle. The chosen reference image is the one which was generated with a viewing direction that best approximates the current viewing direction.

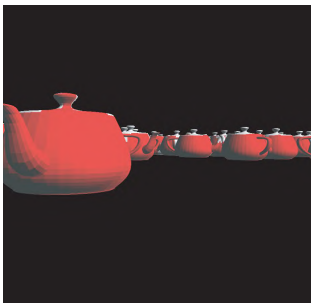


Figure 1. Camera View; occluders are shaded red



Figure 2. RSD; red pixels indicate high DC

During RSD-rendering we have to provide DC values for the subsequent occluder selection. Rholf et al. [1994] define DC as the number of raster operations carried out per pixel. In favor of speed, an approximation is used. Each RSD pixel stores the result from adding up the polygon count of every object whose reference image covers the given RSD pixel. The actual occluder selection is done by evaluating the following equation for each visible object in the RSD:

$$MaxDC_{Obj} \cdot \frac{TotalPolygonCount_{VF}}{VisiblePolygonCount_{RSD}} > \alpha \cdot MaxDC_{RSD}$$

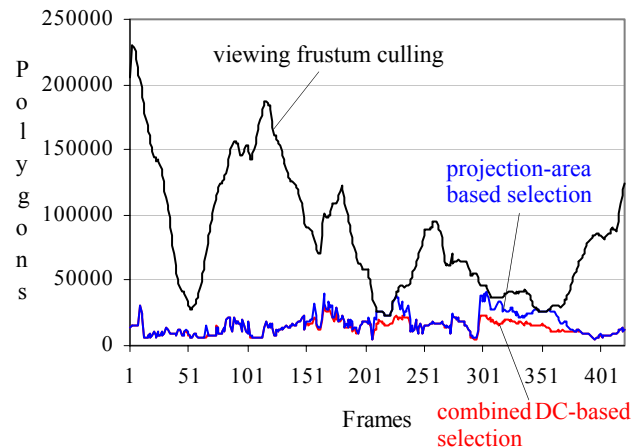
$MaxDC_{Obj}$  refers to the maximum DC value determined by scanning all pixels of the object's visible RSD-area.  $MaxDC_{RSD}$  stores the maximum DC of all RSD pixels.  $MaxDC_{Obj}$  is multiplied by the ratio between the total polygon count of all objects inside the viewing frustum (VF) and the polygon count of all visible objects in the RSD.  $\alpha$  is used to restrict the number of selected objects, in order to keep the polygon count of the occluder-set low.

By multiplying the maximum DC of an object with the estimated ratio between visible and total polygon count, the more objects that are occluded the more objects covering regions of high DC will be selected. Due to the low amount of visible objects in scenes with a high ratio of occluded objects, there is no risk of selecting too many objects for the occluder-set.

If the ratio of visible polygons increases, the number of occluders must be reduced. Otherwise, too many small objects would be selected, that do not necessarily contribute to occlusion when the ratio of occluded objects is low. For this reason,  $\alpha$  is assigned a value between three and four. This has proven useful for a wide variety of scenes; in case of  $\alpha=4$  objects are only selected if the ratio of visible polygons falls below 25%.

The DC based occluder selection approach works in conjunction with conventional size-based occluder selection. The size of the projection area of visible objects in the RSD is estimated by the number of RSD-pixels covered by their corresponding reference images.

The following diagram shows the polygon count to be rendered during a walkthrough of a scene depicted in Figure 1. Occlusion culling is done using the HOM-approach [Zhang et al. 1997], while occluders are chosen by DC based and projection-area based occluder selection. If one of both heuristics becomes true, the object is added to the occluder-set. Combining both selection heuristics results in a noticeable improvement of the culling-rate without raising the occluder-set's polygon count over 6 000 polygons for all frames.



## References

- KURKA, G. 2001. *Bildbasierte Auswahl verdeckender Objekte (image-based occluder selection)*. PhD thesis (205 pages, German), Johannes Kepler University Linz, Austria.
- ZHANG, H., MANOCHA, D., HUDSON, T., HOFF, E. K. 1997. Visibility Culling using Hierarchical Occlusion Maps. In *Proceedings of SIGGRAPH 1997*, Addison Wesley. Whitted, T., Ed., Computer Graphics Proceedings, Annual Conference Series, ACM, 77–88.
- ROHLF, J., HELMAN, J. 1994. IRIS performer: a high performance multiprocessing toolkit for real-time 3D graphics. In *Proceedings of SIGGRAPH 1994*, ACM Press / ACM SIGGRAPH, New York. Glasner, A., Ed., Computer Graphics Proceedings, Annual Conference Series, ACM, 381–395.

# Detail Calibration for Out-of-Core Model Simplification through Interlaced Sampling

Guangzheng Fei\*      Nadia Magnenat-Thalmann  
MIRALab              MIRALab  
University of Geneva      University of Geneva

Kangying Cai              Enhua Wu<sup>†‡</sup>  
<sup>†</sup>Institute of Software      <sup>‡</sup>Faculty of Sci. & Tech.  
Chinese Academy of Sciences      University of Macao

Instead of generating an initial in-core model in first pass and performing a second pass adaptive processing over the original model according to the detail analysis of the in-core model, in this paper we propose a single-pass algorithm to realize detail calibration by introducing a scheme of interlaced sampling in order to obtain higher efficiency.

## 1 Interlaced Sampling

By analogy with the well-studied interlaced sampling techniques in image processing area, which are commonly used to eliminate aliasing, we have found that we can also employ the interlaced sampling method in out-of-core simplification algorithm to achieve adaptive tessellation of the model.

Instead of using a uniform grid as the sampling pattern, we adopt an interlaced pattern formed by the uniform grid and its shifted version, where each grid point is generated by a shift of half a grid size in all three dimensions from its original position. Figure 1 shows a 2D representation of the sampling pattern, where the grid in blue represents the original cell boundaries and the grid in red represents the shifted ones. A uniform clustering method is then performed for both grid settings simultaneously, which results in two individual simplified models at the same time, which will be called the original simplified model and the shifted simplified model, or *Ori-Simp* Model and *Shift-Simp* Model for short respectively.

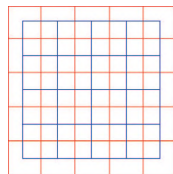


Figure 1: Interlaced Sampling Pattern.

## 2 Computing Degree of Sensitivity (DoS)

We know that uniform clustering methods are sensitive to grid settings. Different grid settings result in different simplified models even in the same grid resolution. It is obvious that places with more detail are more sensitive to the variance of grid placement. In other words, the degree of sensitivity (*DoS*) may indicate how much detail one place has. We define *DoS* for each vertex  $v$  in the *Ori-Simp* Model as  $Q_v(v_{nearest})$ , where  $Q_v$  represents the formula defined by the quadric error matrix associated with  $v$ , and  $v_{nearest}$  represents the nearest vertex to  $v$  in the *Shift-Simp* Model. Since the space has already been partitioned by both grids, to find the nearest vertex in the *Shift-Simp* Model requires very few steps, thus the computation of *DoS* is efficient. Left portion of Figure 2 shows the distribution of the computed *DoS* over the *Ori-Simp* Lucy model, where the *DoS* values have been normalized into range [0, 1].

## 3 Adaptive Merging of the Two Models

Once the *DoS* of all vertices to the interlaced grid pattern is computed, we can then adaptively merge the two individual models

\*email: fei@miralab.unige.ch

based on the distribution of the *DoS* to produce the final output model. Vertices with greater *DoS* are considered having more detail, thus the neighboring regions of these vertices will be densified through further re-triangulation. Regions around the vertices with *DoS* below certain value are kept intact.

The merging of the two models is conducted by two sequential steps, projection and re-triangulation. The projection step is to find the corresponding region in the *Shift-Simp* Model to be merged into the *Ori-Simp* Model. The re-triangulation step is to actually combine the two corresponding regions, taking into account some manifold restrictions. Detail calibration is performed through the combination of the two models. After these two steps, the revised *Ori-Simp* Model is then output as the final simplified model.



Figure 2: Distribution of *DoS* Value and Detail Calibration Result.

## 4 Results and Conclusion

Right portion of Figure 2 shows the resultant calibrated Lucy model, where detail contained in the *Shift-Simp* Model has been adaptively re-populated to the *Ori-Simp* Model to achieve detail calibration. The main contribution of this paper is that a single-pass adaptive scheme is proposed for out-of-core model simplification, where both high quality and high efficiency are maintained. The novel use of *DoS* for detail evaluation has enriched the importance definition for model primitives, thus it may be possible to inspire some new model simplification algorithms.

## References

- FEI, G., CAI, K., GUO, B., AND WU, E. An adaptive sampling scheme for out-of-core simplification. *Computer Graphics Forum, to appear*.
- LINDSTROM, P. 2000. Out-of-core simplification of large polygonal models. In *Proceedings of SIGGRAPH 2000*, Computer Graphics Proceedings, Annual Conference Series, ACM.
- SHAFFER, E., AND GARLAND, M. 2001. Efficient adaptive simplification of massive meshes. In *Proceedings of IEEE Visualization 2001*.

# Digital Pyro for *Reign of Fire*

Patrick Dalton  
The Secret Lab  
patrick.dalton@  
disney.com

Rob Rosenblum  
The Secret Lab  
rob.rosenblum@  
disney.com

Shyh-Chyuan Huang  
The Secret Lab  
chyuan.huang@  
disney.com

Lawrence Lee  
The Secret Lab  
lawrence.lee@  
disney.com

Hank Driskill  
The Secret Lab  
hank.driskill@  
disney.com

Our goal was to create realistic digital pyrotechnic effects for the film “Reign of Fire”. The specific effects we needed were fire, smoke, fog, flying embers and dust kicked up from dragons flying overhead. We also needed the effects to interact with the environment as well as the dragons. To make the challenge more difficult, we realized that it wasn’t good enough to create real world physical interactions: the film needed imagery that was larger than life and exaggerated. We needed to be able to art direct the simulations.

## 1 Simulation

The basic motion of all our effects and their interactions were simulated using computational fluid dynamics (CFD). Generally, phenomena such as smoke and fire have a characteristic flow, but their underlying motions are similar. However, physical simulations are difficult to control, and even after many iterations the results still may not tell the story. Therefore, our workflow had the basic philosophy of allowing viewing and editing of the data at each stage of the pipeline. Our digital pyro effects pipeline had the following stages:

- CFD scene setup and voxelization
- CFD simulation
- CFD simulation viewing and editing
- Particle simulation using CFD results
- Particle viewing and editing
- Rendering

Setup, viewing, and editing were accomplished using custom tools built on top of the Houdini animation package. Our CFD simulator, *smaug*, generated a description of the air flow at each frame of the simulation. Its computational model was based on the work of Nick Foster described in “Modeling the Motion of a Hot, Turbulent Gas” and the work of Jos Stam described in his various papers on modelling gaseous phenomena. *Smaug* used three kinds of data as input:

- Configuration information which specified the regular grid and physical parameters that are used for the simulation.
- A description of the initial conditions for the simulation.
- Animating velocity grids (one per time step) representing moving objects.

Except for the configuration file, input and output data was generated in a Houdini-native data format to facilitate viewing and editing. Each output file contained all the same information as the next frame’s initial condition, letting us edit simulations that have become numerically unstable and restart them at the previously unstable frame using the newly edited conditions.

After generating the flow field, we used Houdini to create particles to visualize the basic look of the effects. For fog, the particles were distributed throughout the flow field. For the others, the particles were generated at descriptive sources in the scene. The quantity of particles varied depending upon the effect. Sometimes secondary motion was explicitly added to the particles to create motion features that were too fine for the CFD simulation to generate.



## 2 Rendering

Smoke, fog, dust and fire were volume rendered using in-house custom shaders written for Steamboat Software’s *Jig* renderer, and the embers were rendered using Houdini’s *Mantra* renderer. The rendering technique varied depending upon the effect.

For embers, a small glowing sphere was instanced onto each particle. For smoke and dust, a volumetric sphere was instanced onto each particle. Each volumetric sample’s opacity was attenuated by a 3D turbulence function to create a wispy, smoke-like feel. For fog, each particle maintained a fixed UVW coordinate, and at each volumetric sample the neighboring particles’ UVW coordinates were blended to create a 3D texture coordinate. The sample’s opacity was then attenuated by a 3D turbulence function evaluated at that texture coordinate.

For fire, we implemented a methodology inspired by PDI’s work in “Shrek”. The particles were used as fields to create an implicit surface, and the space enclosed by the surface was volume rendered. Each particle maintained its own rendering attributes. At a volumetric sample, the particles whose radius of influence contributed to that sample had their rendering attributes blended together, and the total field strength for all contributing particles was summed. If the total field strength fell above a certain threshold value, the sample was considered to be inside the fire, and the blended rendering attributes were used to calculate the final color and opacity of the sample.

The fire state (hot, cooling, smoke) was driven by the temperature value assigned to the particles by the CFD simulation. Sample color was read from a color map, where the color map coordinate was a function of temperature, sample depth within the volume, and field strength. The volume surface was displaced by applying turbulence to the field strength. To accentuate the edges of the displacements, additional color was removed as a function of the displaced volume normal. When the fire state evolved to smoke, the rendering algorithm reverted to the same algorithm used for smoke and dust renders. The raw fire renders were image processed and combined into a fiery look.

## 3 References

1. FOSTER, N. AND METAXAS, D. 1997. Modeling the Motion of a Hot, Turbulent Gas. In *Proceedings of ACM SIGGRAPH 1997*, pp. 181-188.
2. FEDKIW, R., STAM, J., AND JENSEN, H. W. 2001. Visual Simulation of Smoke. In *Proceedings of ACM SIGGRAPH 2001*, pp. 15-22.
3. LAMORLETTE ET AL. 2001. Shrek: The Story Behind The Screen, *SIGGRAPH 2001 Course 19*.



# Diminishing Head-Mounted Display for Shared Augmented Reality

Masayuki Takemura\*  
Yuichi Ohta

University of Tsukuba, Institute of Engineering Mechanics and Systems, Tsukuba, 305-8573, Japan

## 1 Background and problem

Eye-contact plays a very important role in the face-to-face human communication. In the shared augmented reality space, however, the eye-contact between two people is blocked by the head-mounted displays. An HMD is a necessary device for merging the virtual and real worlds to realize the augmented/mixed reality. [Ohta and Tamura 1999] As a side effect, it covers the user eyes. In a shared AR/MR space, multiple users want to share the real world as well as the virtual objects. However, the eye-contact information in the real world is dropped because of the HMDs.

## 2 Key idea

The user's face hidden by an HMD is synthesized as a virtual object by using an image-based rendering technique and it is overlaid to the real user's face to diminish the HMD. The user's line-of-sight is measured by a detector embedded in the HMD and it is represented on the synthesized virtual face. Then his partner can recognize his line-of-sight.

## 3 System

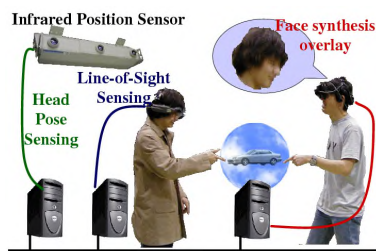


Figure 1: Diminishing HMD

Figure 1 illustrates the conceptual overview of the system to diminish HMDs. An infrared position sensor is used for measuring the head-pose of the users. Users wear video see-through HMDs with line-of-sight sensing capability. The head-pose and line-of-sight data are supplied to the graphics PC and the user's face is synthesized and overlaid to the view of real world in video rate.

## 4 Face synthesis

The user's face with eye expression is synthesized from his/her several photographs. The face synthesis process has two stages, offline stage and online stage. Figure 2 illustrate the outline of face synthesis.

**Offline stage:** Face photographs with three different poses are prepared and the feature points on the face are manually corresponded. The 3D shape of face is reconstructed by using a computer vision technique. An expanded texture map is generated from the three photographs. The eyeball texture is made as a mosaic of several eyeball photographs. Two sets of 3D model and expanded texture are prepared for faces with open eye and closed eye.

\*e-mail: takemura@image.esys.tsukuba.ac.jp

### Online stage:

This stage is performed in video rate. The 3D model for arbitrary eyelid status is generated by the morphing technique. The motion of line-of-sight is represented by a translation of eyeball texture. The 3D model is located at the real head position of a user and the virtual camera is located at the real camera attached to another user's HMD. A PC with graphics accelerator is enough for the real time rendering. Figure 3 shows some snapshots of user's view.

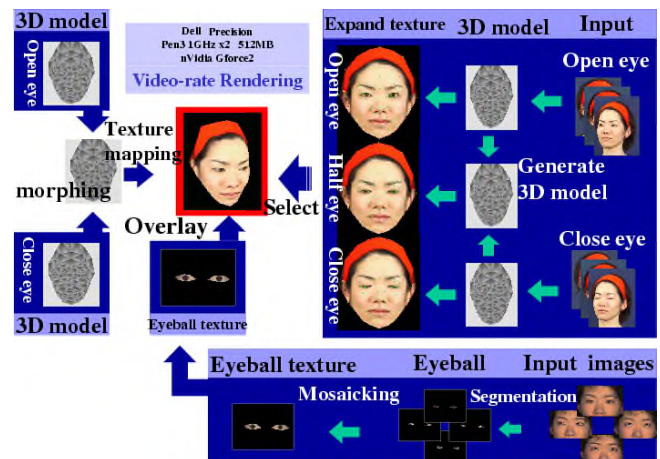


Figure 2: Face synthesis from photographs



Figure 3: Snapshots of user's view

## 5 Future direction

In the present implementation, the whole face is synthesized and overlaid because of the difficulty to merge the partial virtual face with the real face seamlessly. Photometric registration should be investigated as well as geometric registration.

## References

Ohta, Y. and Tamura, H. Eds: Mixed Reality, Merging real and virtual worlds, Springer-Verlag, 1999.

**Demonstration Video:** <http://www.image.esys.tsukuba.ac.jp/~takemura/LAB/study/siggraph/index.html>

# Diorama Engine - A 3D Video Storyboard Editor for 3D Computer Animation

Koji Mikami  
Tokyo University of Technology  
koji@so.teu.ac.jp

Toru Tokuhara  
Keio University  
torukun@sfc.keio.ac.jp

Mitsuru Kaneko  
Tokyo University of Technology  
mits@media.teu.ac.jp

## 1. Outline

In this sketch, we will demonstrate a video storyboard tool targeted for 3D computer animation. Our tool provides limited but specialized functions of standard 3D computer graphics software, focusing on ease of scene construction, camera control and the ability to preview in realtime. This allows for quicker and easier creation of video storyboards over existing approaches.

## 2. Background

Creating video storyboards using 3D computer graphics tools should be convenient because information such as scene data and camera paths can form a foundation for work in the production stage. However there are issues involved.

- (a) A 3D computer graphics tool and hardware is required
- (b) Time is required for learning the tool and creating the picture

Creating 3D video storyboards will require special staff and hardware additions. To make practical use of the 3D video storyboard, we believe that a specialized tool that is easy to use and low on hardware requirements is needed.

## 3. Prototype Development

The Diorama Engine is being developed as a specialized tool for creating video storyboards. The current implementation has functionality on staging (figure 1) and shooting (figure 2).

The Diorama Engine is able to accomplish the following compared to standard 3D computer graphics software.

- (a) Easy enough to use for inexperienced 3D users or users who are not comfortable with the PC
- (b) Low hardware requirements (consumer desktop computers or notebook computers is sufficient)
- (c) Realtime preview allows trial and error and results to be instantly previewed
- (d) The document file is compact in size, so remote collaborations are practical

## 4. Evaluation

To evaluate the functionality of the tool, the implemented prototype was user-tested as an exercise activity by 100 students studying motion picture production at School of Media Science, Tokyo University of Technology. In spite of their first experience, most students were able to operate the tool without problems after a 20 minute lecture on operating instructions, and some were able to finish a video storyboard within the following 45 minute exercise. Also, the tool ran without noticeable slowdowns on a notebook computer.

## 5. Future Work

To improve functionality as a video storyboard tool, the ability to handle dialogue, sound, and music is being added. Also, cut editing functionality is being implemented. With these additions, an extended use of the tool as a motion picture directing simulator can be expected.

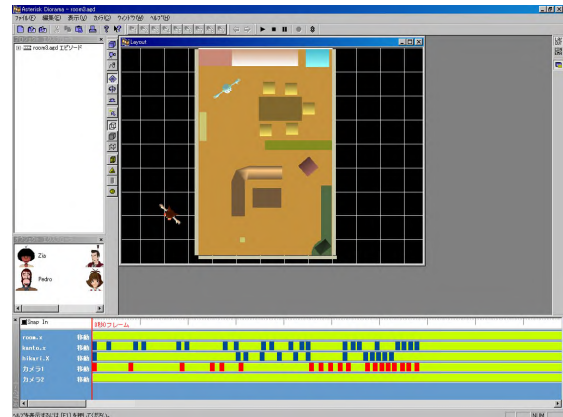


Figure 1. Staging in Diorama Engine

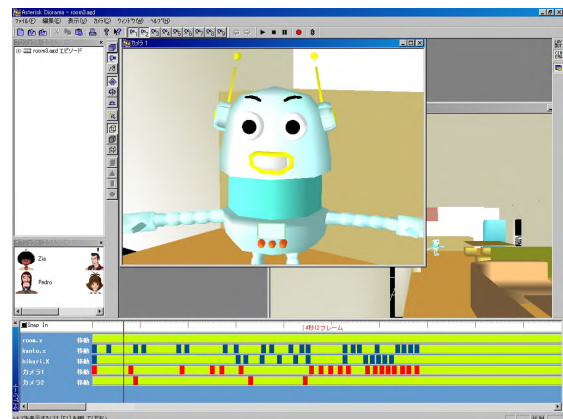


Figure 2. Shooting in Diorama Engine



# DIVIPRO: Distributed Interactive Virtual PROtotyping

Mashhuda Glencross\*, James Marsh, Jon Cook, Sylvain Daubrenet, Steve Pettifer and Roger Hubbard  
Advanced Interfaces Group, Department of Computer Science,  
University of Manchester, UK

## 1 Introduction

DIVIPRO is a prototype tool for simulating the assembly and disassembly of mechanical engineering components. It is aimed particularly at situations where the responsibility for decision making is shared between geographically dispersed design teams, and provides a collaborative environment in which different team members concurrently visualize and manipulate models. It is currently being evaluated using models from the aerospace and medical industries.

DIVIPRO includes a number of components required to work together in real time: a geometric kernel for importing STEP models and performing geometric queries; a choice of two geometric constraint engines (basic and advanced depending on the level of functionality required); a flexible body simulator for modeling deformation of objects such as plastic tubes and cable harnesses; and a haptic module which allows forces from collisions, constraints and flexible components to be experienced by the user through a PHANToM force-feedback device.

The DIVIPRO distribution architecture aims to maximize user responsiveness while maintaining consistency, and it is this aspect of the project that forms the basis of this sketch.

## 2 The DIVIPRO System

Typically distribution in commercial CAD systems is performed using technologies such as Microsoft's NetMeeting, which aim to synchronize the displays of multiple users. This prevents users having different viewpoints, or simultaneously interacting with models. Each user in the DIVIPRO system has a different viewpoint and may simultaneously interact with models. This interaction is mediated by a constraint engine, the job of which is to check for collisions and apply constrained motion.

Distribution is managed by utilizing the University of Manchester's DEVA framework [Pettifer et al. 2000]. To a first approximation DEVA has a client-server architecture. This simplifies the issues of synchronization since the definitive state of an entity is only stored in one location. The disadvantage of this approach is that when a user interacts with a model; a round trip from the client to the server has to be made in order to validate the interaction, checking for collisions and constraints, and to update the state of the entity. If the server and clients are all located on a fast network (eg. a company's internal LAN) this round trip may be acceptable. However, with DIVIPRO we are targeting geographically dispersed design teams - a trial DIVIPRO session involved users located in Germany, Spain and the UK connected together via the Internet.

Due to the large distances involved (there is a 400ms ping time from the UK to Spain) the lag experienced by the users begins to make a client-server architecture unusable. In contrast a peer-to-peer architecture which performs the collisions and constraints calculations locally at each client maximizes responsiveness but at the cost of increased complexity in synchronizing the multiple clients.

Rather than being a single process, however, the DEVA server consists of a number of separate processes, which may be distributed across the network. Conceptually the server processes are

presented as a single logical server, with their physical location hidden from clients. At any one time an entity is managed by a single server process, however it can migrate between server processes effectively "moving" across the network. This migration process is hidden from clients via an automatically updated naming layer.

DIVIPRO employs this novel migration technique to move the constraint engine to a location on the network closest (in terms of latency) to the user that initiates an interaction with a model. This minimizes the round trip time required to validate an interaction, resulting in a faster update rate both for this user and other observers.

In practice the overhead of migrating the constraint engine entity is too great since the complex data structures required to efficiently detect collisions and constraints are time consuming to rebuild. This problem was overcome by having multiple constraint engines simultaneously present, and migrating a simple token which permits only one of the constraint engines to be active at any one time. The inactive constraint engines listen to validated updates from the active one, thereby maintaining synchronization.

If a second user simultaneously begins to interact with the model, then that user is forced to use an already active constraint engine, which may not be optimal. The architecture minimizes the response time for a single user interacting with a model, eg. in a scenario that involves turn taking, but also maintains consistency for multiple interacting users, such as in a shared assembly task. Other migration strategies are currently being investigated: if multiple users are interacting then a fairer solution may be to migrate the constraint engine entity to a location equidistant from the interacting clients, or possibly to allow multiple constraint engines to be simultaneously active provided the objects the users are manipulating are sufficiently distant from each other that they do not interact.

The authors would like to thank the European Commission for providing funding and the following members of the DIVIPRO consortium: Labein, Sener, KL Technik, DLR, Eurocopter, and the University of Salford.

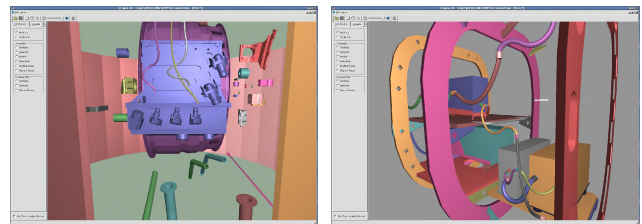


Figure 1: Screen shots showing the DIVIPRO system being used to assemble complex models.

Further information about DIVIPRO and other research projects at Manchester can be found at <http://aig.cs.man.ac.uk/research/>.

## References

- PETTIFER, S., COOK, J., MARSH, J., AND WEST, A. 2000. Deva3: Architecture for a large scale virtual reality system. In *Proceedings of ACM Symposium in Virtual Reality Software and Technology*, ACM Press, 33–39.

\*e-mail: khotem@cs.man.ac.uk

# DocuDrama Conversations

Leonie Schäfer  
Fraunhofer Institute FIT  
leonie.schaefer@fit.fraunhofer.de

Elaine M. Raybourn  
Sandia National Laboratories  
emraybo@sandia.gov

Amanda Oldroyd  
BTexact Technologies  
amanda.oldroyd@bt.com

## Introduction

This sketch describes DocuDrama, a tool that offers a generation of interactive narratives that are based on activities in a collaborative virtual environment. DocuDrama [Schäfer et al. 2001] is built as part of TOWER [2002], a Theatre of Work Enabling Relationships, which allows project members to be aware of project relevant activities as well as to establish and maintain the social relationships that intensify team coherence.

## Challenge

DocuDrama aims at recording the history of desktop events and activities generated by a project team in a collaborative work environment. The replay of the team's interactions with documents is visualised by avatars that enact the events as they occurred in the shared workspace.

The challenges in generating narratives of project histories are twofold. On the one hand the challenge is to select significant events and to combine event sequences to new, meaningful representations. On the other hand the challenge is to present the project's history and progress in an entertaining way which captures the users' attention and conveys complex information fast and effectively.

## Approach

In our approach the history replay takes place in the TOWER multi-user, virtual environment. TOWER offers several virtual landscapes that visualise shared workspaces and their documents. Event sensors monitor the users' activities in the project's shared workspaces, which are then played out by avatars as symbolic actions. In this approach the history replay on DocuDrama focuses on the interaction between project members on documents. The resulting DocuDrama conversation can exist between two users interacting at approximately the same time, or if one user interacts with a document that has been created by another user.

Present and past activities take place in TOWER world in the same environment. While the avatars of the present are displayed in colour, the avatars of the past are visualised as ghosts in black and white, and small in size. The ghost avatars replay event sequences that indicate interactions, e.g. the joint editing of a paper or collaboration on documents towards a project deadline. The DocuDrama conversation uses elements of non-verbal communication [McCloud 1993], e.g. the avatars turn towards each other before performing a symbolic action.

The figure shows a sequence of a DocuDrama conversation. In front two ghost avatars are interacting with each other on top of a document, while in the background avatars enact ongoing activities in the project's shared workspace.

## Conclusions and Future Work

DocuDrama provides an efficient and entertaining way of presenting summaries of project-relevant events and activities.

We performed three sets of discount usability studies on DocuDrama. Our findings to date indicated that a successful DocuDrama story should include rich contextual information about the project, people, and progress. Additionally, more analyses will be necessary as we experiment further with emotions and displaying affect in our stories.

In the near future we plan improve the conversational behaviour of the avatars. We will also enhance the storyline of the history replays, e.g. with story-marks symbolizing the subject of interaction. Finally we are investigating to improve the camera control as a major instrument in generating the DocuDrama films.

## Acknowledgements

We thank the members of the TOWER project team for their contribution to DocuDrama. The TOWER project is partly funded through the IST program IST-10846.

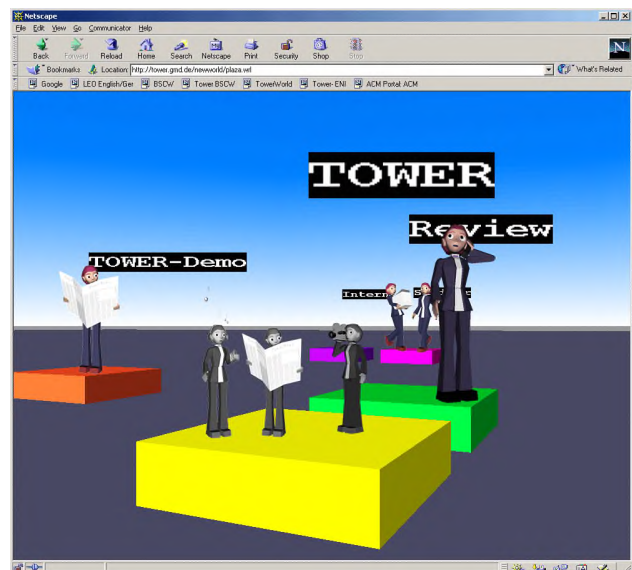


Figure 1. Scene from a DocuDrama Conversation

## References

- SCHÄFER, L., PRINZ, W., OLDROYD, A., GAVIN, L., 2001. *Virtual Storytelling of Cooperative Activities in a Theatre of Work*. In *Proceedings of the International Conference ICVS 2001: Virtual Storytelling*, Balet, O.; Subsol, G.; Torguet, P. (ed.), Berlin: Springer, 191-200.
- TOWER PROJECT, 2002. <http://tower.gmd.de/>
- MCLOUD, S. 1993. *Understanding Comics*. Kitchen Sink Press, Northampton, MA.

# Dragon Scales: The Evolution of ScaleTool for *Reign of Fire*

Ernest J Petti  
Walt Disney Feature Animation

Thomas V Thompson II  
Walt Disney Feature Animation

Adolph Lusinsky  
The Secret Lab

Hank Driskill  
The Secret Lab

## 1 Introduction

In recent years there have been several movies starring creatures with scaled surfaces. Among these are *Jurassic Park*, *Dragonheart*, and *Lake Placid*. The surfaces of these creatures have generally been constructed by layering painted textures atop displacement maps. This gives the model texture, but the scales stretch and shrink under the movement of the creature, giving a rubbery look that is not realistic.

ScaleTool is an extension of FurTool, a proprietary tool used by The Secret Lab in *Dinosaur* and *102 Dalmatians*. FurTool grows fur randomly across NURBS surfaces on a per patch basis, hiding patch boundaries through random, yet density preserving, growth. For ScaleTool we were tasked with designing a system that grew scales in an ordered pattern. Since these scales are geometry, and not textures, size is maintained under model deformation.

## 2 Look Development

It is the job of Look Development to groom the scales on the hero model. The layout and look of the scales is broken into three components: pose, growth, and material.

### 2.1 Pose

The pose of the scale is described by length, width, bend, curl, and twist attributes. Since a model may contain thousands of scales, it is impractical to modify each scale individually. Therefore, ScaleTool provides the notion of "key scales", a uniform grid of scales placed on each patch. The user can modify any attribute with the grown scales deriving values from neighboring keys. For finer control, an attribute map can replace the key attributes.

Upon completion of the ideal groom, some amount of noise must usually be applied to give the creature a more natural look. ScaleTool provides several noise attributes and compensators. For example, when noise is applied to width the groomed spacing between scales is maintained. Since the scale can be laid flat in any orientation, noise in the length, width, and twist must all be considered when maintaining the groomed spacing during growth.

### 2.2 Growth

The growth algorithm walks each patch placing scales along a row until the edge of the patch is reached. This process continues with the next row until the patch is completed. The initial step size in each direction is based on the total number of scales to be placed.

Unfortunately, adjacent patches can be parameterized or shaped differently, making patch boundaries visible. Much of our work on ScaleTool targeted removing this visual discontinuity. First, a correction compensates for patch shape by multiplying each step size by the ratio of edge arc lengths. Next, to deal with non-uniform parameterization, ScaleTool generates density compensation maps, storing values of  $\delta P/\delta u$  and  $\delta P/\delta v$  from sampled points on the surface. The growth algorithm queries these maps to adjust the step size. If the parameterization changes significantly between steps, sub-step measurements of the maps are taken. The final result is a uniform pattern of scales on each patch with no noticeable patch boundaries, using an algorithm that did not require any knowledge of adjacent patches or calculations of distances from one scale to the next (Figure 1).



Figure 1: Final image of the dragon.

### 2.3 Material

In order to provide a convincingly realistic look, RenderMan surface and displacement shaders are applied to each of the scales. However, the look of the scales is not determined in isolation from the underlying surface. Rather, important parameters of the scales and underlying surface are automatically passed into the shaders where they can modulate the color or displacement. This technique allows patterns, such as stripes that span a series of scales, to be pulled from the surface onto the scales (Figure 2).

## 3 Collision Avoidance

Since the dragon is an animated model, the deformation of the model could result in scales either penetrating or flaying. ScaleTool preserves the original groom across frames by trying to maintain the distance from each scale to the model's surface. Each scale has one end attached to the model and the other end floating. It is this floating endpoint that must be checked. We first find the "virtual attachment" for this point using Direct Parametric Tracing (a haptic rendering algorithm) in combination with bisection. If the distance of the endpoint is different for the current model than for the reference, then that scale is rotated toward a space point along the surface normal that is the original distance away. This algorithm runs fast enough that even on the dragon model, with thousands of scales, the overhead is minimal as only seconds are added to a render time of hours.

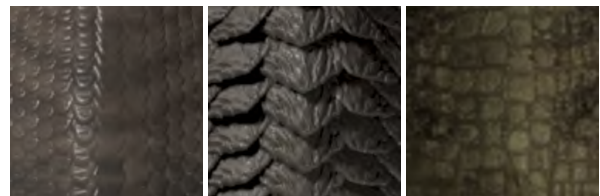


Figure 2: Close-up of scales generated for (a) snake, (b) early dragon, and (c) final dragon.

<sup>1</sup>{Ernest.Petti, Tom.Thompson, Adolph.Lusinsky, Hank.Driskill}@disney.com



# Drawing With Feeling: Designing Tactile Display for Pen

Ivan Poupyrev<sup>1</sup>, Shigeaki Maruyama<sup>2</sup>

<sup>1</sup> Interaction Lab, Sony CSL  
poup@csl.sony.co.jp

<sup>2</sup> Sony Semiconductor Network Company  
shigeaki.maruyama@jp.sony.com

## 1 Introduction

There is a sense of satisfaction using a pen or pencil to write or draw. Feeling the imperfections on the paper as the pen tip moves over it and observing marks emerge create the inherently physical and intimate feeling of drawing. Our extensive discussions with artists and designers suggest that this physicality and intimacy with drawing not only brings enjoyment, but also perhaps assist in the artist's creative process.

Unfortunately, most of this physicality is lost when computers are substituted for paper. Today's tablet-based pen input devices provide only one channel of communication – from the user to the computer. Hence, the user not only loses the natural properties of real drawing, but furthermore, the potential benefit of using a computer is not fully exploited to aid the creative process.

We aim to overcome these problems by enhancing the pen with a miniature tactile display to allow users to interact with the drawing in the most natural way – through his or her hand. Our tactile pen is based on the TouchEngine™, a novel actuator, which is miniature, lightweight and low-powered, making it easy to embed even into small devices such as the pen. We are investigating interactive applications of pen tactile interfaces in a number of application scenarios described below. Unlike related projects [1], we do not attempt to simulate exactly the physical properties of the pen and paper. Instead, drawing with the computer should, perhaps, feel different, supporting functions unique to drawing with the computer and not possible with a real pen and paper.

## 2 TouchEngine™ tactile display

The major obstacle in developing a tactile display for a pen is the actuator. We do not know of any actuators that have low latency, are small enough to fit inside a pen, are low power to allow wireless construction, and can provide enough physical force all at the same time.

The TouchEngine is an actuator that we designed to meet these requirements. It is constructed by sandwiching thin piezoceramic film (0.25  $\mu$ m) in between adhesive electrodes, which makes a thin ~0.5mm beam (Fig. 1a). The piezoceramic material works as a solid state “muscle” by either shrinking or expanding depending on the polarity of the applied voltage. The material on the top has an opposite polarity than the bottom, so when a signal is applied, the entire structure bends. Hence, this configuration is often called

the *bending motor*. Bending motors that currently exist consist of only two layers (biomorphs) and require a minimum of  $\pm 40$ V, making them unsuitable for small devices. By sandwiching multiple layers of piezo material, while keeping total thickness the same, we can significantly reduce the required voltage and produce sufficient feeling with only 5-7V.

Two strategies have been developed to produce tactile feeling with the TouchEngine. In *direct tactile display*, the actuator would move at an area of the pen where the user touches (fig 1, b). In *indirect tactile display*, the actuator is placed somewhere inside the device with a weight attached. When the actuator bends, the mass moves rapidly up or down moving the entire device with equal momentum in the opposite direction. This allows the user to feel a force impulse from the pen.

## 3 Drawing with tactile feedback

To investigate how tactile feedback can enhance drawing, we embedded a TouchEngine into a Wacom pen and used it with a Sony VAIO LX PC, which has the Wacom pen input technology built into the screen. Although we currently use wire to connect the pen to the driving circuit, a wireless version is under construction.

We investigated a number of scenarios. In one scenario, the user is able to feel patterns, lines, or colors allowing the user to distinguish parts of an illustration by touch. Because the Wacom tablet tracks the pen even when it does not touch the tablet, in another scenario, the user is able to feel control points for Bezier curves before selecting them making it much easier to pick the desired control point. We also experimented with the use of tactile feedback during free-hand drawing and allowing the user to feel different colors. By combining feedback with pressure sensitivity, the user can also feel the line thickness.

The initial demonstrations received positive response from artists and designers. We believe that enhancing drawing with tactile feedback will open new and exciting opportunities in interactive computer graphics applications. We are currently developing a wireless version of this device, enhancing existing application scenarios, and investigating new applications.

## References

1. Baxter, B., et al. DAB: Interactive Haptic Painting with 3D Virtual Brushes, *Proceedings of SIGGRAPH 2001*, pp 461-468

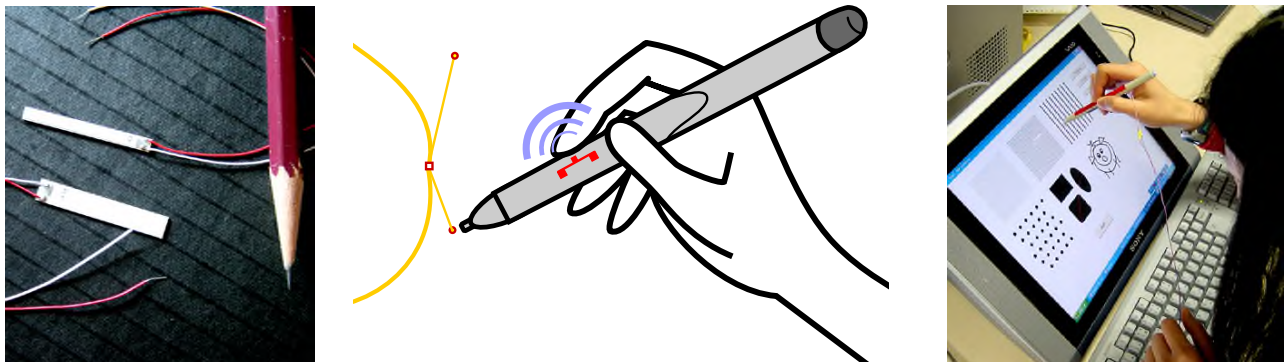


Figure 1: (a) Actuators: 3 mm and 5 mm width, 0.5 mm thick (b) pen construction (c) initial applications

# Dynamic Simulation of Wing Motion on "Reign of Fire"

Carlos Gonzalez-Ochoa

David Eberle

Rob Dressel

Disney Feature Animation\*

## 1 Introduction

The challenge of animating believable dragons was presented to Disney Feature Animation and The Secret Lab (TSL) for the film "Reign of Fire". The film called for 100ft creatures with wing spans of 300ft that could undergo enormous speeds and accelerations. The artistic direction required each dragon to have wings that transition between a variety of physical behaviors and interact with the environment. To solve this challenge the Muscle and Skin system used in Disney's Dinosaur [Eskuri 2000] was extended with a variety of controls to do physical simulation. In this sketch we discuss some of the issues encompassing the creation of this simulator and give an overview of the successful and unsuccessful paths taken during its development.

The character consisted of several layers of muscle and skin. The muscle layer was driven by an animated rigged bone structure. An intermediate layer, called the inner skin, was bound to the muscles of the body but not the wings. The wing membranes were physically simulated and indirectly controlled by the bones, providing a separate inner skin layer. Finally an outer skin was driven by the combination of these two inner layers. This layered approach enabled the TDs to have a high-resolution seamless outer skin that combined the motion from our simulator with that of the muscles of the body.

## 2 Controlling Dynamics

The dynamics were controlled by an elaborate simulation system. At the beginning of the production the simulator started as a simple mass-spring system. Although the system was general enough for other applications of dynamics such as soft bodies or cloth; it became obvious that the animated motion of the rig did not work well with the physical assumptions of this system. As a result, several techniques were developed to accommodate the physical constraints of specific shots. When the creature was subjected to high speeds during flying, unwanted stretching typically occurred. To maintain the material look by increasing the spring forces caused the simulator to take many small steps at a great computation cost. To solve this problem with more efficiency we introduced a post-dynamic relaxation technique based on [Provot 1995].

Another technique devised to deal with this issue was a way to decouple the dynamics from the translation of the rig. Instead of using the total motion from the rig to drive the simulation, the motion of a separate coordinate frame could be used to control the translational behavior of the system.

Directing dynamics to obtain particular shapes at specific times has always been a challenge for animators. We added blending functionality to allow animators to guide the dynamics by using sculpted surfaces as targets. One method interpolated between the target shape and the calculated dynamic surface. The second method biased the dynamics to move towards the target shape. In either case a keyable attribute controlled the amount of blending between the two surfaces.



Figure 1: Dynamic Wings

## 3 Collision Detection

Since the character had to react with the environment and itself, a suite of collision primitives were provided. Solids and in particular height maps were optimized for the crawling scenes over rough terrain. The problem of self-collision proved to be the most challenging. Common animation motion resulted in constant wing self-intersections, sliding contacts, and high velocity collisions. Sampling issues were alleviated with a temporal triangle-triangle collision test described in [Provot 1997].

Frequently persistent collisions would occur in areas that visually could be ignored and did not contribute any significant physical behavior. Attribute maps were employed to identify such regions to avoid collision response calculations.

## 4 Conclusion

The purpose of creating a dynamic simulation is to simplify the work of the animators. If the simulator's only inputs are initial conditions and animation curves, then the time saved in animation is easily consumed by technical directors fighting the system to get the desired look. The challenges brought by this production pushed the development of a dynamics system that provided secondary motion, but more importantly allow user control. We believe that providing robust control methods are essential to resolve production struggles between simulation and animation.

## 5 Acknowledgements

We would like to thank Philip Schneider for his efforts and contributions to the system.

## References

- ESKURI, N. 2000. The art and technology of disney's dinosaur. ACM Press, ACM Siggraph 2000 Course Notes, ACM Siggraph.
- PROVOT, X. 1995. Deformation constraints in a mass-spring model to describe rigid cloth behavior. *Graphics Interface*, 147–154.
- PROVOT, X. 1997. Collision and self-collision handling in cloth model dedicated to design garments. *Graphics Interface*, 177–189.

\*cgo@acm.org, david.eberle@disney.com, rob.dressel@disney.com



# Dynamic Skin Deformation and Animation Controls

## Using Maya Cloth for Facial Animation

Jimmy Chim  
MFA Computer Art  
School of Visual Arts  
jchim@acm.org

Hyunsuk Kim  
MFA Computer Art  
School of Visual Arts  
Hyunnat@aol.com

This sketch presents a solution for creating expressive computer facial animation using off-the-shelf software. We will describe how to apply dynamic skin deformation on a CG character using Maya Cloth and how to create artist-friendly animation controls. Historically, facial animation with sophisticated skin deformation based on physical characteristics of skin could only be done using proprietary software in big studios, which are not accessible to small studios, independent animators and students. The solution that we have developed is easy to understand and does not involve any programming knowledge.

### Dynamic Skin Deformation

Skin is composed of three layers: the outermost skin layer, the middle sliding soft tissue layer and the bottom muscle layer. As a unit, skin tissue exhibits viscoelastic property, which behaves as both an elastic solid and a viscous liquid.<sup>1</sup> To simulate skin tissue, previous research works suggest the use of a mass-spring model.<sup>2</sup> These methods begin with converting the vertices of a CG model into point mass nodes and the edges into springs connecting them. When a force is applied on a point mass node, it will be passed through the springs to the rest of the model. As a result, the skin can be deformed dynamically.

Blend Shape (also known as morphing) is the most popular tool for creating facial animation. It allows animators to blend and mix various facial expressions together and create a smooth transition between them. Although it provides animators with precise control on skin deformation, the resulting animation lacks the elasticity and dynamic behaviors of skin tissue.

Maya Cloth from Alias|Wavefront is a very powerful tool in generating realistic movements of fabrics and garments. We found that Maya Cloth can also be used to deform skin dynamically while preserving the flexibility and elasticity of skin tissue. By converting a CG head model into a cloth object and applying a set of constraints and animation controls, it is possible to achieve a more organic skin deformation as compared to Blend Shape.

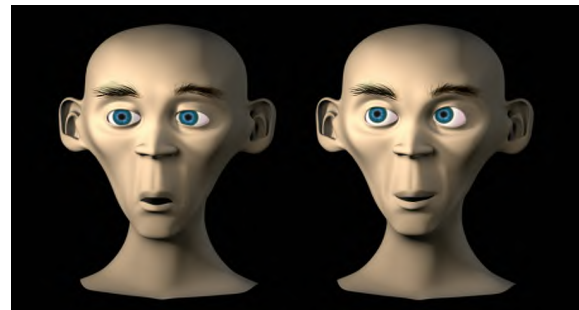
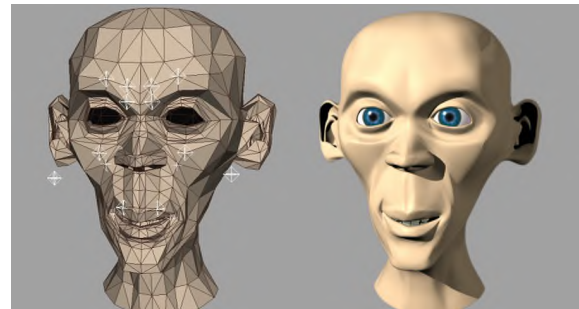
Common to all simulation algorithms, this method takes time to compute the result. We will describe how to reduce the simulation time to an acceptable value. In our demonstration, the character (with 1688 vertices) takes approximately 1.8 seconds per frame to simulate on a AthlonXP 1800+ machine.

### Facial Animation Controls

Maya Cloth enables us to apply dynamic skin deformation, but it may not be easy to animate as it does not provide real-time feedback of skin deformation. To address this problem, we need a set of animation controls to deform the skin interactively. Once we are satisfied with the animation, we can then apply the dynamic skin deformation on the CG character.

We have developed a set of facial animation controls in Maya utilizing either IK Spline skeletons or geometry constraints with Wire Deformers. As compared to regular FK skeleton setup, our facial animation controls enable animators to sculpt the path for muscle deformation around the mouth area. This setup works well for animating the Orbicularis Oris, that has three distinctive movements, stretching, extruding, and moving against the teeth. With our facial animation controls, animators are able to imitate dynamic 3D motions without writing any script. This facial animation controls complement with dynamic skin deformations for creating facial animation.

Being animators, our ultimate goal is to achieve elaborate motions and explicit emotions on CG character to communicate with audiences. However, creating expressive and yet convincing facial animation is a great challenge for animators with limited resources and tools. Our solution can help animators bring CG characters to life.



### References

1. PARKE, F. I., AND WATERS, K. 1996. *Computer Facial Animation*. A. K. Peters.
2. LARRABEE, W. 1986. A finite element model of skin deformation I: Biomechanics of skin and soft tissue: A review. *Laryngoscope*, vol. 96, 399-405

# Dynamics and Dodos: Rigging and Animation Methods for *Ice Age*

Adam Burr  
Technical Director  
Blue Sky Studios

Ross Scroble  
Technical Animator  
Blue Sky Studios

This sketch explains how the Animation Team at Blue Sky Studios executed a challenging sequence for "Ice Age" and describes in detail a particularly useful animation software tool called *followThrough*.

## Pre-Planning and Execution

The Dodo Bird sequence involved many more characters in some scenes than for much of the rest of the movie. We had to figure out how a large number of dodos could be executed in a single shot. We ended up breaking up shots into sub shots, and assembling them in lighting. The animation software we use could handle approximately a dozen dodos before the scene became too slow.

For the more complex shots, where there was a large deal of repeatable animation, we used a proxy system to place the animated characters and match the rate that they were moving. The proxies were placed by the animators and then handed off to the technical directors as a separate layer. As the layers were finished in animation, the images were composited together in Shake™ to get a feel of how everything was working together.

## A Unique Pre Visualization

The visualization of this sequence was a bit unfocused when it arrived in animation. This gave the animators a chance to lend some storytelling and composition to the sequence. Unlike other sequences, three key animators blocked out most of the sequence, created shots, and worked with frame lengths of shots as well as the rough action. These blocked out shots were later dispersed to other animators to be completed.

## Application of Animation Fundamentals

The Animation Department was very excited to tackle the dodo animation. The Dodo's design and character lent to a very snappy, fun style of animation. Prior to the sequence, the animators attended an "Animation Fundamental Workshop" directed by William Frake, an industry veteran who is employed at Blue Sky in the Story and Layout Department. Mr. Frake gave us a jump-start towards a very fundamental approach, which gave the dodos a very 'cartoony' feel in their 3D environment.

## Tools to Assist in Animation

Our proprietary referencing interface and other in-house tools were used to manage large numbers of the same character. Search and Replace functions were added to our selection tools and character toolboxes. Animation import/export software and pose storage tools helped animators achieve scene-to-scene consistency. The time saved in pre-planning was used in achieving the performance of the characters, which was the key to making this sequence a success.

## Dynamics

Besides enhancements to character management software, another tool that quickly became critical to animation was the *followThrough* tool - an animator controlled character dynamics solution. This tool was designed so that the system could be set up and controlled by animators without the assistance of additional technical staff. This was done partly because of staffing limitations and also because the animators did not want to loose control over the characters' performances.

The *followThrough* tool takes a signal processing approach to motion dynamics. The result is a smooth and stable dynamics system that does not require sub-frame time steps. The animator is given control over the rest position of the system and the dynamic movement decays continuously towards this position. The results can be seen clearly in the mammoth's tail. Even with a significant amount of dynamic movement, the tail still hits the animator's carefully choreographed poses.

Ice Age



Frame from final composite showing many dumb dodos



Frame showing *followThrough* on feathers

# Editable Dynamic Textures

Gianfranco Doretto\*  
UCLA Computer Science Department

Stefano Soatto†  
UCLA Computer Science Department

## Introduction

This technical sketch presents a simple and efficient algorithm for editing realistic sequences of images of dynamic scenes that exhibit some form of temporal regularity. Such scenes include flowing water, steam, smoke, flames, foliage of trees in wind, crowds, dense traffic flow etc. We call this kind of scenes *dynamic textures*.

Traditionally the problem is approached either by *physics-based* (PHB) techniques [Popović et al. 2000] or by *image-based* (IMB) techniques [Schödl et al. 2000]. Here we propose an IMB technique, based on the work of [Soatto et al. 2001], that relies on a statistical model of the visual signal. We say that a sequence of images  $\{I(t)\}_{t=1\dots\tau} \in \mathbb{R}^m$  represents a *linear dynamic texture* if it is a realization of a second-order stationary stochastic process. If we call  $y(t) = I(t) + w(t)$  a sequence of images corrupted by white, zero-mean Gaussian noise  $\{w(t)\} \in \mathbb{R}^m; w(t) \sim \mathcal{N}(0, Q)$ , the assumption of second-order stationarity of  $\{y(t)\}$  corresponds to the existence of a positive integer  $n$ , a process  $\{x(t)\} \in \mathbb{R}^n$  with initial condition  $x_0 \in \mathbb{R}^n$  and symmetric positive-definite matrices  $Q \in \mathbb{R}^{n \times n}$  and  $R \in \mathbb{R}^{m \times m}$  such that

$$\begin{cases} x(t+1) = Ax(t) + v(t) & x(0) = x_0; v(t) \sim \mathcal{N}(0, Q) \\ y(t) = Cx(t) + w(t) & w(t) \sim \mathcal{N}(0, R) \end{cases} \quad (1)$$

with  $I(t) = Cx(t)$ , for some matrices  $A \in \mathbb{R}^{n \times n}$  and  $C \in \mathbb{R}^{m \times n}$ .



## System Overview

Referring to the figure above, the overall system takes two types of input and produces one output. The inputs are (a) a finite video clip and (b) a set of editing parameters. The output of the system is a potentially infinitely long video of a dynamic texture similar to the original one, that can be interactively modified by editing the parameters (b).

The system is composed of several “modules.” The first module, *learning*, takes as input a finite, noisy sequence of images  $\{y(1), \dots, y(\tau)\}$  and returns an estimate of the model parameters  $\hat{A}, \hat{C}, \hat{Q}, \hat{R}$ . This is borrowed from the prior work of [Soatto et al. 2001].

Once a model has been learned, new sequences can be trivially generated by the *synthesis* module which simulates the model (1). This entails choosing an initial condition  $\hat{x}_0$  (for instance one of the original images), drawing a sample of an IID Gaussian process with covariance  $\hat{Q}$ , performing one step of the iteration  $\hat{x}(t+1) = \hat{A}\hat{x}(t) + \hat{v}(t)$  and computing the new synthesized image as  $\hat{I}(t) = \hat{C}\hat{x}(t)$ . However, one would like to insert an *editing* module before the synthesis. This is described in the next section.

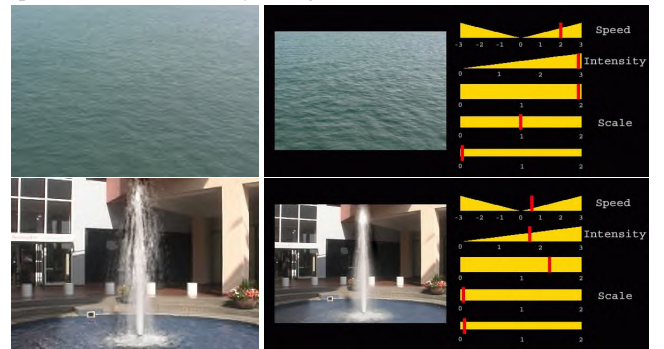
## Editing

This module allows the editor to (1) simulate a novel sequence that has the same spatio-temporal statistics as the input video clip, and

(2) interactively modify the spatio-temporal characteristics of the simulation in order to achieve the desired perceptual effect. In principle, any modification of the system parameters, for instance  $\hat{A}, \hat{C}, \hat{Q}$ , results in a novel synthetic sequence whose spatio-temporal statistics is altered with respect to the original sequence. Unfortunately, naive manipulation of the model parameters rarely results in sequences that have any resemblance with reality. Therefore, this module allows manipulating groups of parameters while enforcing causality, stability, and other constraints that are necessary to yield a realistic outcome.

The sketch shows how to change the spatial scales by deforming the actual sub-space spanned by the columns of  $C$ . This is done by substituting the matrix  $\hat{C}$  with the matrix  $\tilde{C} \doteq \hat{C}W$ , where  $W \in \mathbb{R}^{n \times n}$  is a diagonal matrix with the non-negative real numbers  $w_1, \dots, w_n$  as its diagonal entries. To alter the speed of a synthesized video clip, we proved that the frequencies associated with the poles of the system matrix  $A$  have to be multiplied by the same constant factor  $\Omega$ . Therefore, each normalized frequency  $\psi_i$ ,  $i = 1, \dots, n$ , is replaced by  $\tilde{\psi}_i \doteq \Omega\psi_i$ . In order to reverse the time axis, so as to invert the visual flow of a given dynamic texture, while maintaining the same temporal dynamics (speed), all we need to do is to substitute  $\hat{A} = V\Lambda V^{-1}$  with  $\tilde{A} = V^*\Lambda V^{-1}$ , where  $\Lambda$  and  $V$  respectively are the diagonal matrix of eigenvalues and the matrix whose columns are the corresponding eigenvectors of  $\hat{A}$ , and  $*$  denotes the complex conjugate. Finally, notice that for a non-zero  $Q$ , the input  $v(t)$  excites the modes of the state  $x(t)$  and causes it to evolve as a discrete-time Brownian motion. The intensity of the driving noise determines how far away from the initial condition the Brownian motion travels. We show experiments where we rescale the noise components by a positive constant  $\alpha$ , i.e. we substitute  $\hat{Q}$  with  $\tilde{Q} \doteq \alpha\hat{Q}$ .

The sketch shows examples of dynamic textures like “Ocean,” “Fountain,” “Fire,” and “Smoke,” and their sped up, slowed down, and time reversed versions, incorporated with effects induced by spatial scale and intensity changes.



## References

- POPOVIĆ, J., SEITZ, S. M., ERDMANN, M., POPOVIĆ, Z., AND WITKIN, A. 2000. Interactive manipulation of rigid body simulations. In *Proceedings of SIGGRAPH 2000*, 209–218.
- SCHÖDL, A., SZELISKI, R., SALESIN, D. H., AND ESSA, I. 2000. Video textures. In *Proceedings of SIGGRAPH 2000*, 489–498.
- SOATTO, S., DORETTO, G., AND WU, Y. 2001. Dynamic textures. In *Proceedings of IEEE International Conference on Computer Vision*, vol. 2, 439–446.

\*e-mail: doretto@cs.ucla.edu

†e-mail: soatto@ucla.edu



# Embodied Interaction

Embodied Interaction, explores the components of interactive media, with a primary focus on the role that interaction design has on virtual experiences. My electronic artwork is concerned with the transformation of the human species, specifically its biological components and its behavioral characteristics. This transformation or evolution is an environmental reaction to the manifestations of science and technology. This presentation examines the need to address the physical body and how the action of users needs to be interconnected with the interface and content of a interactive piece. From the development of opaque sculptural input devices to the use of transparent technologies my interactive installations seeks to examine the process of conditioning users; their predetermined interaction; and the physicality of computing.

The actual conduit of electronic content, or to use the buzzword interactive multimedia, is made up of three interconnected components: (1) the information structure, the organizational space, (2) the interaction, what users are actually doing with their physical bodies, (3) the interface which sits between the body and information.

As an electronic media artist, I am particularly concerned with the structure of interactivity and methods of developing a deeper connection between the content and its conduit. This has led me to examine the relationship between the interface and the interaction. I have become increasingly concerned with the physical body's role in virtual environments.



## Opaque Interaction

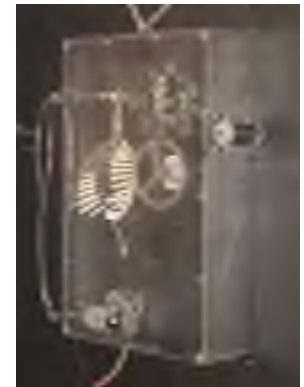
By using sculptural input devices, the viewer is forced to physically behave in manner interconnected with the projected virtual interface. This may constitute a more passive interaction, by confining them to a chair with simple buttons in the armrest which correspond to binary choices within a nonlinear narrative. Or be a laborious activity of moving a cumbersome device which controls the navigation of a virtual environment. Both use an opaque interaction, where the user is consciously using their bodies in a manner which has a one to one relationship with the activity on the screen.

Bill Hill

Jacksonville University  
2800 University Blvd North  
Jacksonville, FL 32211  
whill@ju.edu

## Transparent Interaction

Conversely, with transparent interaction, there is a more concealed, less apparent relationship between the user and the output. Instead of a conscious focus on their own bodies and the reaction of the interface, here they react to the interface. By using sensors and video tracking these machine sculptures respond to the activity of the audience. *Black Lung* is a artificial lung machine which responds to the physical activity in front of it. With increased movement the single lung inflates and deflates at a faster pace. The title comes from Black Lung disease, common among coal miners who excavate energy from the earth to power the machines of technology and are physically weakened to the point where they must rely on a machine for life.



*Statue* is an interactive kinetic video installation, where a single monitor rotates in a 360° circle allowing the viewers to view a virtual sculpture. Using a video tracking system, the monitor rotates to follow the position of the audience. A 3D generated human figure of classic proportion is displayed on the screen and rotates proportionally with the monitor. The rendering of the human figure transforms between a solid opaque character and a transparent illustration of its anatomy. The concept beneath this work is the social ideal. The notion of socially directed body modification dates back thousands of years; the Greek "super-anatomical" sculptures helped to invent the ideal form, the 19th century development of moving pictures illustrated the behavior and movement of that form, but it wasn't until the technological development of "X-Rays" that the real notion of the body changed.

In either method of interaction design, when we use computers, we are affected by them, we take away a residue of the experience with us - in our body. The body is our interface with the world. It is the point of contact with the content.

## Introduction

High-resolution voxel rendering is no longer the pipe dream it once was. In the age of multi-gigabyte RAM, voxels are finding an ever-increasing role in creating cutting-edge visual effects. Digital Domain has developed a unique, animator-friendly way of working with voxels that is revolutionizing the way volumetric effects are done.

## The Evolution

We began work on our voxel package three years ago as a way to model and render a large, static storm system for a simulation ride. A new approach was needed because traditional ray marching or particle rendering would have been impossibly slow. The resultant idea was to use 'cloud points' to fill a voxel buffer, then save a series of texture maps - pre-lit cross sections of the cloud - and render them on cards. In this way, a whole fly-through animation could be done without having to re-calculate volumetric lighting for every frame - the whole cloud was pre-baked into a set of texture maps. A shader added fractal noise to the generally soft maps during rendering, adding detail.

The interface for our voxel editor/renderer was a Side Effects Software's Houdini SOP (Surface Operator) - this allowed the modeling of our voxel cloud to blend seamlessly with our workflow. The user could drag cloud points around in Houdini, and immediately see a series of semi-transparent OpenGL grids showing the volumetrically shaded voxel results in the viewport. This gave the user a degree of feedback unheard of by any other volumetric technique.

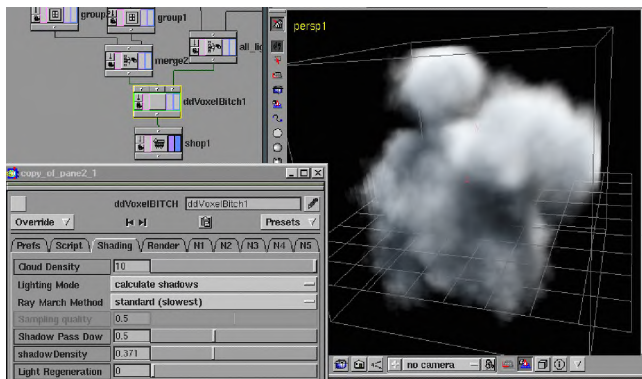


Figure 1. The VoxelB interface, created as a Houdini SOP, gives the user real-time feedback via OpenGL.

As other productions at Digital Domain began to see the usefulness of the tool, they began to utilize the voxel tool for a variety of new purposes it had not been designed for. New features and options were added as creative demands grew, until the interface became a mess of options and parameters. A more generalized approach to voxel modeling and animation was needed, and voxel compositing was born.

## New Approach

A scripting editor was incorporated into the SOP to break the voxel modeling process into simple operators, which could be chained together to create a complex result. Operating much like a 2D compositor, the newly dubbed "VoxelB" allowed unprecedented manipulation of voxel information, including

blurring, sharpening, clamps, ramps, distortion (using vector voxel buffers), math functions, cloud points, voxelization of polygonal geometry, noise layering, and merging of voxel buffers. There was almost nothing that couldn't be modeled and animated using this approach.

A mantra of resolution-independence was adopted, which assured that all voxel operations were based on real-world units and positions, never the size of a voxel or the resolution of the buffer. This assured that voxel work done in low resolution would match the results when resolutions were increased. If a user changed the bounds of a voxel buffer, instead of stretching and translating the contents it merely 'captured' a different region of a larger, implicit volumetric environment.

The usefulness of VoxelB has far exceeded our initial expectations and intent. Almost every project that comes to DD comes up with new and often unexpected uses for this tool. Apart from the obvious cloud, smoke and dust effects, voxel modeling by this approach has allowed us to create realistic fire, churning white-water, avalanches, time-lapse rock erosion, melting glaciers, growing trees and bushes. Although every project pushes VoxelB into new areas of use, its core has remained essentially unchanged - a testament to the flexibility of its procedural approach. Its renderings have been featured in *Dr. Seuss' The Grinch That Stole Christmas*, *Red Planet*, *X-Men*, *Lord of the Rings*, and *The Time Machine* as well as numerous commercials.



Figure 2. Voxel water horses in "Lord of the Rings"



Figure 3. This eroding arch from "The Time Machine" was modeled and animated using VoxelB and turned into geometry using Houdini's ISO sop.



# Example-Based Interpolation of Human Motion

Hiroshi Mori  
University of Tsukuba  
hmori@edu.esys.tsukuba.ac.jp

Jun'ichi Hoshino  
University of Tsukuba / PRESTO, JST  
jhoshino@esys.tsukuba.ac.jp

## Abstract

Generating realistic human motions from sparse pose descriptions are important for computer animation and virtual human applications. In this sketch, we propose an example-based motion interpolation technique using independent component analysis (ICA). The proposed method is also useful for human motion conversion, motion blending, and converting symbolic descriptions to continuous motion.

**CR Categories:** I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism - Animation; G.3 [Mathematics of Computing]: Probability and Statistics – Time series analysis; E.4 [Data]: Coding and Information Theory – Data compaction and compression; J.5 [Computer Applications]: Arts and Humanities – Performing Arts

**Keywords:** animation, ICA, character behavior.

## 1 Independent Component Analysis for Human Motion

We extract a few independent motion signals from multi-dimensional human movement. We used FastICA[1] because it showed the stable performances in our experiments. The result of the extracted motion signal corresponds to the perceptually meaningful motion features that characterize the complex human actions. By modifying these motion trajectories in ICA domain, various motion characteristics can be produced.

The important characteristic of ICA is that the base vectors construct the motion feature space. Interpolation of the arbitrary points in the base vector space is possible. We can obtain the plausible motion by interpolating arbitrary sparse pose in ICA space.

## 2 Motion Editing using ICA

Our technique can provide a flexible motion editing capabilities like a traditional keyframe animation, yet reuse of motion capture data is possible at various levels. The implementation of the algorithm is easy because the publicly available ICA softwares can be used for our purpose.

Based on the motion interpolation using ICA, we have developed several basic techniques including:

- 1) Motion texture mapping: The detail of the subtle motion expressions can be reconstructed as small variations in the action trajectories in ICA domain. We search the similar motion structures in samples, and add residuals to the interpolated motion.
- 2) Motion conversion: We can convert the motion characteristics by manipulating motion trajectory in ICA domain. For example, we can convert motion signals to a particular person's motion.
- 3) Blending concurrent behavior: Blending concurrent motions are possible by applying ICA hierarchically.
- 4) Generating motion from symbolic descriptions: We can reconstruct motion with fewer breakdowns from symbolic motion descriptions.

Another important application of our technique may be intelligent storyboard system[2]. By interpolating sparse pose descriptions,

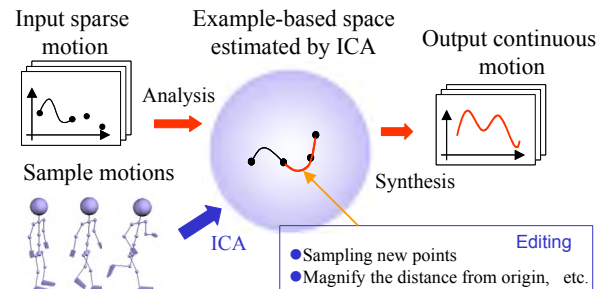


Figure1: Overview of example-based interpolation of human motion

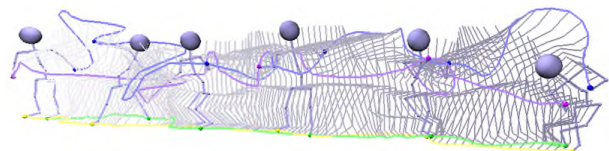


Figure 2: Example of generating a dance sequence from six sparse poses. The complex motion between the specified pose can be recovered by using example-based interpolation technique.

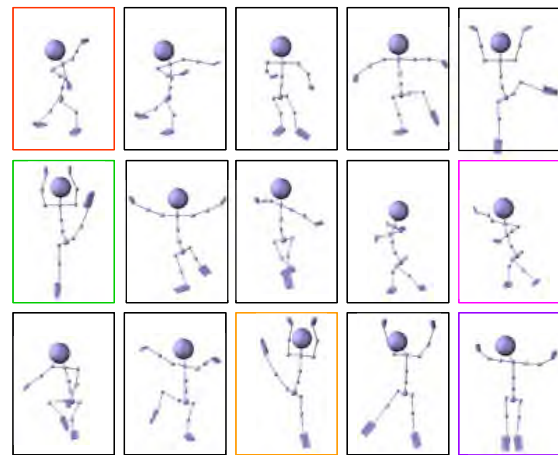


Figure 3: Example of generating a ballet sequence. We can obtain the plausible and expressive motion compared with simple keyframe interpolation.

we would be able to generate continuous motion from sparse pose specifications.

## References

- [1] A.HYVARINEN: Fast and Robust Fixed-Point Algorithms for Independent Component Analysis , IEEE Trans. on Neural Networks, 1999
- [2] JUNICHI HOSHINO, YUMI HOSHINO: Intelligent Storyboard for Prototyping Animation, IEEE Int. Conf on Multimedia and Expo, ICME2001, Conference CD-ROM FAI.03 ,2001

## Exploiting Screen Space

This sketch concerns the development of a renderer which takes advantage of the two dimensional nature of the frames it produces. The project is chiefly inspired by the way in which artists have used 2D abstraction in the past, from scribbled cartoons to cubist paintings. It would seem that exploiting the plane of the canvas can help us to make more interesting compositions.

The aim is to create a system that renders 3D geometry by applying vector based drawing tools to each frame. This would allow an artist to make animated films that look hand shaded without the tedium of physically coloring every frame. Much excellent work has already been done in the area of drawing object *edges* in screen space, but the majority of sketchy *shading* algorithms produce images that betray the 3D forms of the underlying geometry. Post-production filters that distort images in two dimensions can yield stylish stills, but their simplicity is too often revealed when applied to a sequence of frames.

I propose a system that gathers shading information from the 3D scene, and then draws directly on to the frame using the kind of tools found in 2D image editing packages. These tools could be anything from simple brush strokes, to polygon shapes or stamps, their actions described as vector information for storage and manipulation. For the most part the tools correspond to those used in traditional artistic disciplines (pens, paint brushes), though tools that do not exist in the real world could just as easily be programmed.

A number of styles have been implemented so far as stand-alone C++ programs. Their results are displayed here applied to Phong shaded spheres lit by a single spot light source. Figure 1 shows the geometry rendered by the Maya renderer.

In figure 2 an airbrush tool draws strokes around object edges based on the light direction. Specular shading is rendered by airbrush dots. This gives objects a soft, flat look that could not easily be attained from a conventional renderer.

Figure 3 is a colored marker sketch simulation. Contiguous areas of similar luminance are crosshatched by colored splines over blocks of color. These splines move smoothly from frame to frame achieving a continuity that would be laborious for a sketch artist to maintain.

Figure 4 demonstrates the possibility for more abstract rendering styles. Objects are approximated as semi-transparent polygons that are merged with the original image. Edges are drawn by wide, blotchy brush strokes.

Currently, the program inputs bitmap images rendered from an animation package, but the desired system would work directly from geometry files. This would give the program access to information on occlusions and object movement, allowing more calculated decisions about placing and distorting strokes. It would also be more accurate than a bitmap-based method, and therefore capable of processing small or distant geometry.

Ian Mackinnon, ianmackinnon@hotmail.com  
NCCA, Bournemouth University  
Fern Barrow, Poole, BH12 5BB, UK

Depth of field and motion blur are conceivable additions, along with further deformation tools, specifically, distorting the 3D geometry of a scene based on manipulation of vector data in screen space. The final touch would be adding an interactive GUI with which non-technical users could design custom brushes and shading algorithms.

In conclusion I believe that the quality, variety and distinctiveness of the images produced so far are sufficient to warrant further work into 2D vector rendering systems.

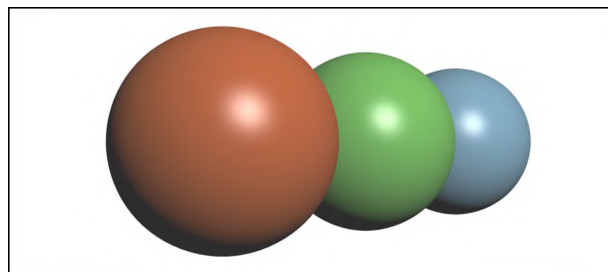


Figure 1

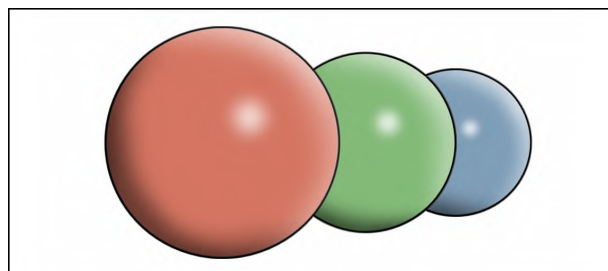


Figure 2

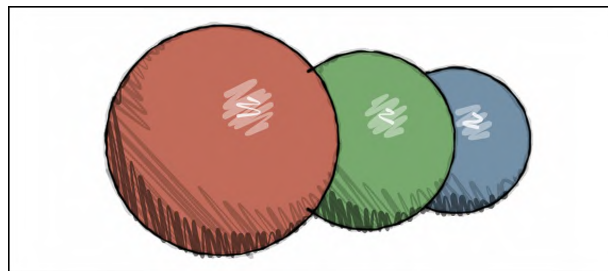


Figure 3

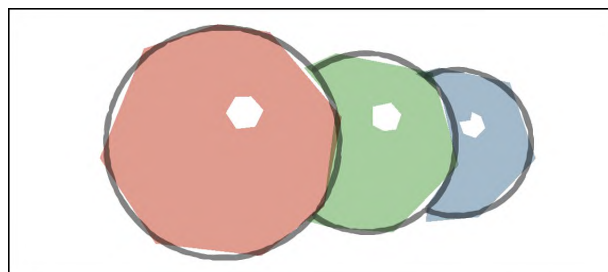


Figure 4

# Expressive Features for Movement Exaggeration

James W. Davis\* Vignesh S. Kannappan  
Dept. of Computer and Information Science  
Ohio State University

## 1 Introduction

Given a single motion-capture sequence of a person performing a dynamic activity at a particular *intensity* (or effort), our goal is to automatically warp that movement into a natural-looking exaggerated version of that action. Consider warping a movement of a person lifting a lightweight box to make the movement appear as if the box were actually very heavy. We describe an efficient data-driven approach applicable to animation re-use that learns the underlying regularity in an action to select the most “expressive” features to exaggerate. Other “style-based” approaches are presented in [Gleicher 1998; Brand and Hertzmann 2000; Vasilescu 2001].

We first collect motion-capture examples of a person performing an action at different intensities (See carrying example in Fig. 1). The data are placed into a matrix  $A$  with each column corresponding to a specific DOF trajectory, with each intensity set following the other. Singular Value Decomposition,  $A = U\Sigma V^T$ , is used to generate a representative trajectory basis  $U$  from which we perform efficient computations to rank trajectory expressibility.

## 2 Identification of Expressive Trajectories

The “expressiveness” of a joint-angle trajectory across different intensities is based on how large and how predictable the changes appear. *Observability* measures the magnitude of change across intensity, and is calculated as the Euclidean distance of trajectory projections on  $U$  between the first and last intensities. We additionally weight the influence of this change with the sum of bone lengths of the joint’s children. To determine if those changes have *predictability* (to enable correct exaggeration), we compute an eigenvalue-weighted sum of projection correlations across intensity on each basis vector.

The expressiveness  $E$  for a joint-angle trajectory across intensity is computed as  $E = O \cdot P^2$ , which effectively attenuates the observability values when the predictability is low. The expressibility plot for the carrying examples is presented in Fig. 1. The resulting expressibility values for all the joints are normalized, sorted in descending order, and accumulated until reaching a desired percentage of expressiveness. This results in a smaller set of key features.

## 3 Exaggeration Models

We model the selected trajectory projections across intensity with intensity-to-coefficient mappings. Given a new motion-capture sequence, the selected expressive trajectories are extracted and projected onto  $U$ . Assuming an arbitrary starting intensity index for this motion, we use the intensity-to-coefficient mappings to produce new projection coefficients that either consistently exaggerate or diminish the current intensity of the motion. The approach is easily extended to include multiple people in the training set. We union the top expressive features for each person and average the intensity-to-coefficient mappings to create a generalized model.

\*e-mail: jwdavis@cis.ohio-state.edu

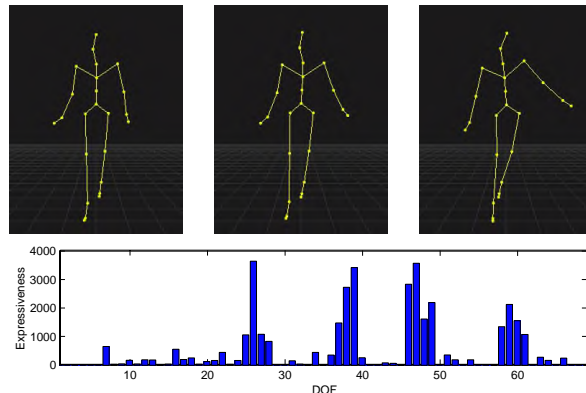


Figure 1: Top: Motion-capture sequences of a person carrying a bag (not shown) with weight 0, 25, and 50 lbs (shown left to right). Bottom: Modal expressiveness plot for the carrying examples.

## 4 Experiments

Using a Vicon-8 motion-capture system, we collected light, medium, and heavy carrying sequences for two people of different height, weight, and gender. A single walk cycle was automatically extracted and time-normalized to 50 samples. The expressive feature set and intensity-to-coefficient mappings were computed.

To validate the selected expressive features, we swapped those key trajectories between the lightest and heaviest examples for one person. The results were nearly identical to the actual motions. To examine the capability of the averaged model, we warped person-1: light-to-heavy and person-2: heavy-to-light. The resulting motions appeared natural. We additionally generated novel exaggerations for a person not used in training. Within limitations, the model exaggerated the intensity to produce believable movements.

## 5 Conclusion

We presented a fast and efficient method to exaggerate motion-capture with naturalistic results. The approach identifies the most expressive motion features within a low-dimensional sub-space and models the regularity across performance intensity. The resulting system can exaggerate new motion sequences without the need for additional supportive input.

## References

- BRAND, M., AND HERTZMANN, A. 2000. Style machines. In *SIGGRAPH 00 Conference Proceedings*, 183–192.
- GLEICHER, M. 1998. Retargetting motion to new characters. In *SIGGRAPH 98 Conference Proceedings*, 33–42.
- VASILESCU, M. A. O. 2001. Human motion signatures for character animation. In *SIGGRAPH 01 Conference Abstracts and Applications*, 200.

# Firefighter Training Virtual Environment

Tazama U. St. Julien, Chris D. Shaw

{stjulien,cdshaw}@cc.gatech.edu

GVU Center, Georgia Tech, 801 Atlantic Dr, Atlanta, GA, 30332

The Firefighter Training Simulation is a virtual environment being developed at Georgia Tech in collaboration with the Atlanta Fire Department. The VE user is a commanding officer trainee who instructs teams of virtual firefighters to perform different actions to help put out virtual fires. This simulation was developed using the Simple Virtual Environment (SVE) Library, an extensible framework for building VE applications [Kessler 2000].

The VE includes a furnished single family home, a fire truck, a fire hydrant, various firefighting tools, and firefighters. To put out the fire, the user issues low-level task commands to the virtual firefighters of the form "team# task destination". For example, typing "1 chop 2" will instruct the fire team 1 to chop at location 2. The firemen execute keyframed animations to display their assigned task being carried out. The command entry to the system is performed by an operator who translates the verbal command from the trainee user into the appropriate syntax for the system.

We created a small motion database of firefighter character animations that includes cutting with a chainsaw in various positions, chopping with an ax, walking, crawling, climbing a ladder, carrying various tools, carrying an injured person, connecting and pulling the hose, and spraying the hose. We are currently using path finding and collision detection algorithms to find paths for the firefighters and avoid collisions with walls of the house.

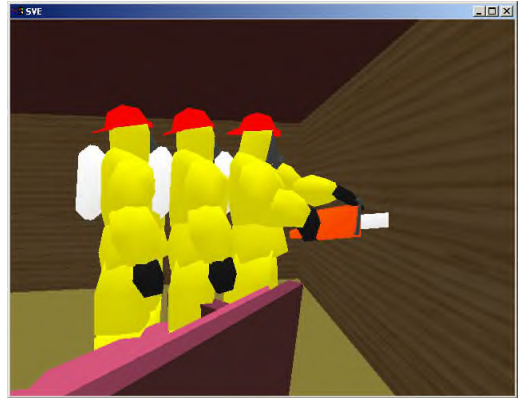


## Fire Simulation

We are using NIST's Fire Dynamic Simulator (FDS) to compute realistic physical fire and smoke behavior. FDS is a computational fluid dynamics (CFD) model of fire-driven fluid flow. It numerically solves the Navier-Stokes equations appropriate for low speed, thermally driven flow with an emphasis on smoke and heat transport from fires. We model the house offline with FDS and extract fire and smoke data for the entire house volume at one-second increments. We use this pre-computed data to visualize and animate the fire and smoke in our simulation.

To draw fire and smoke, we are implementing a voxel-based splatting renderer, which uses the FDS smoke density to control the opacity and color of each screen-aligned splat. Our splat-based Doppler radar [Jang et al. 2002] renderer runs in real-time, so we

are confident that this smoke and fire rendering will also yield real-time update.



For interactive simulation, we must have data available for every sequence of actions that the firefighters can perform. When a firefighter sprays water on the fire, the blaze should subside and more smoke/steam should be produced. Therefore we must use FDS to simulate every possible scenario. We create a decision tree, which loads the appropriate FDS data corresponding to the sequence of the actions performed. We will limit the amount of actions being performed in order to keep the size of the tree and the amount of data needed manageable.

To minimize storage requirements for volumetric fire and smoke playback, we have developed a volume compression technique that quantizes the smoke density, temperature, and heat values to 8 bits per voxel, and applies hyperdimensional prediction and differencing to achieve lossy compression rates of 90 to 95%.

## Conclusion

In collaboration with the Atlanta Fire Dept, we have developed a prototype VE simulator for the training or evaluation of trainee fire commanders. The AFD is enthusiastic about the current prototype, and plans to start using it by the Fall of 2002.

## References

- D. KESSLER, D. BOWMAN, L. HODGES 2000. The Simple Virtual Environment Library: An Extensible Framework for Building VE Applications, *Presence: Teleoperators and Virtual Environments*. 9(2): pp. 187-208.
- JANG, J., W RIBARSKY, CD SHAW, N. FAUST 2002. View-Dependent Multiresolution Splatting of Non-Uniform Data. In *Proceedings of IEEE Visualization Symposium*, Barcelona.



# Folded: Negotiating the Space between Real and Virtual Worlds

Dr. Samantha Krukowski / the Folded Team  
Convergent Media

[samantha@rasa.net](mailto:samantha@rasa.net) / [www.folded.org/team.html](http://www.folded.org/team.html)

## Abstract

A painting is a landscape of possibilities (fig.1) – a form field, a material experiment, a background or foreground, a place of play and imagination. Photographs of people looking at paintings reveal them looking into them, finding things in them that were never consciously put down, never put *on* the canvas. The way people look at paintings requires a painter to remember the world outside of the work that it, itself, points to. Not only is the materiality of the painted field interesting--its image, vibrant hues and liquid surfaces. What is increasingly interesting, given the possibilities of expanded visualization systems, is how a painting might move into and out of itself, towards and away from its material existence, based on what may or may not be found in its home medium.

**CR Categories:** I.3.7 Three-Dimensional Graphics and Realism, I.6 Simulation and Modeling, K.m Miscellaneous and K.8.0 Games.

**Keywords:** Interdisciplinarity, Immersive Environments, Kinetics, CAVEs, Interactivity, Media, Virtual Reality.

## 1 Overview

*Worldmaking as we know it always starts from worlds already on hand; the making is a remaking.* — Nelson Goodman, Ways of Worldmaking

The images in the painting of origination are inspired by a host of sources, some scientific, others cosmological, some historical, others material, some digital, others derived from the act of painting itself. It is valuable to find images through the paint, to utilize the postures of painting to make the next form. Images in paintings do not have stable sources. Working between the material and ephemeral worlds of painting, there are fragments, transparencies, mutations of form, color, dimension and sequence.

Considering and reconsidering a painting is an experiential, rather than a methodological, set of actions. If a painting is the primary plane or surface from which to derive an immersive environment, the order and components of that painting are set free. We can extend the life of the kinetics of the visual. A segment of a painting can be subjected to a series of changes that not only alter the schema of the original but that extend and transcend its character over time. Overlays of the mobility and transparencies expose new image sets, image sets that also move and change. When projected, these image sets may be utilized as the ground of a new painting, allowed to run and change and interact with the more permanent marks they inspire.

At the heart of an exercise like this is the understanding that the relationship between digital imagery and material imagery is grounded in equal exchange — lest we lose ourselves completely in the belief that digital images are the only images of the future, there is value in the wonder and exploration of how the digital and non-digital act in the in-between.



Fig.1: a landscape of possibilities



Fig. 2: a painting as immersive environment



Fig. 3a: a sculpture or building translated from a note

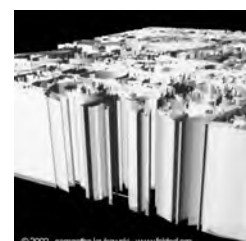


Fig 3b: a city extruded from the painting

## References

- FOUCAULT, M., on Velazquez' *Las Meninas*. 1966 in *Les Mots et les choses*, Paris.
- SARTRE, J-P., 1948. *La recherche de l'absolu*.
- CAGE, J., 1969 *Notations*, New York: Various Works
- SWISS Minimalism, by HERZOG, DEMEURON AND ZUMTHOR
- SAPER C., 2001. *Networked Art*.
- DOLLENS, D. 2001. *D2A Digital to Analog*.
- KRUEGER M., 1991. *Artificial Reality II*.
- KOOLHAAS R., 1995. *Delirious New York*. New York.
- LYNN G., 1993. *Folding in Architecture* in *Architecture Design Profile No 102*.
- DELEUZE, G., 1988. *Le Pli Leibniz et Le Baroque*.
- CACHE, B. 1995. *Earth Moves*, MIT Press.



# FlowField: Investigating the Semantics of Caress

Timothy Chen, Sidney Fels  
Electrical & Computer Engineering  
University of British Columbia  
Vancouver, B.C., Canada  
{tichen, ssfels}@ece.ubc.ca

Thecla Schiphorst  
Interactive Arts  
Technical University of British Columbia  
Surrey, B.C., Canada  
schiphorst@techbc.ca

## 1. Introduction

We have created a new interactive experience piece called FlowField. Participants touch and caress a multi-point touchpad, the MTC Express, in a CAVE (CAVE Automatic Virtual Environment), directly controlling a flowing particle field. Collisions in the particle field emit musical sounds providing a new type of musical interface that uses a dynamic flow process for its underlying musical structure. The particle flow field circles around the participant in a cylindrical path. Obstructions formed by whole hand input disturb the flow field like a hand in water. The interaction has very low latency and a fast frame rate, providing a visceral, dynamic experience. In FlowField, participants explore interaction through caress, suggesting reconnection with a sense of play, and experiencing a world through touch.

## 2. Interaction Experience

The FlowField installation allows participants to immerse themselves in a virtual flow field of particles (Fig. 1) displayed on a four-screen CAVE. By using hands and fingers on the MTC Express, users can interact directly with the flow by introducing obstructions in the path of the particles (Fig. 2). The effect is not unlike placing one's fingers into a stream, experimenting with blocking the flow of water. The experience is complemented by a dynamic soundscape caused by the particles striking the obstructions. The combination of sound and visual sensations brings the interaction activity to life, ebbing and flowing in intensity with the particle flow, as directed by you.

By applying one's hands and fingers onto the MTC Express in different ways, interesting flow patterns can be developed that are aesthetically pleasing. Multiple touchpads can be used to enable a group of users to interact with the same flow field, creating an interplay of particle and sound between users, who can work with, independently, or against each other to generate complex interference patterns, all while immersed in the swirl of activity.



**Figure 1:** Cylindrical particle field. In virtual space, users are standing inside the cylinder while the particles revolve around them.

## 3. Motivation

The human sense of touch is as well developed as our senses of sight and hearing. Yet, while there are plenty of devices that deliver and receive visual and auditory information, the same cannot be said for the tactile sense. Even with very complex and compelling virtual reality display systems such as the CAVE, there is often little means of providing tactile input. The FlowField system instead uses the pressure-sensitive, multi-point MTC Express, developed by Tactex Controls, Inc. to give users a new form of tactile input.

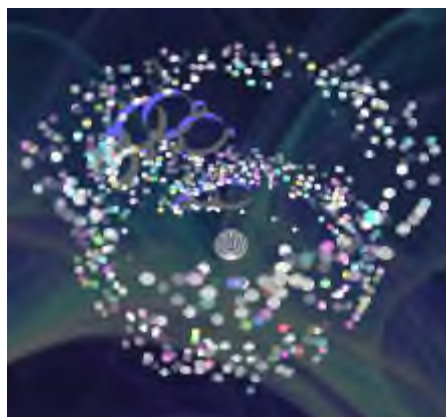
Another purpose of this work is to explore the semantics of gesture and its application in interactive systems. In particular, the nature of the touchpad allows us to study interactions between the hand and a solid surface. The important property of this type of interaction is the inherent repeatability of gestures when performed on a fixed surface. This is in contrast to unconstrained gesturing, such as sign language, which can be recognized using glove-based devices or video capture.

## 4. Related Work

FlowField exhibits several attributes necessary for a compelling interactive installation, including instantaneous response, implemented with low latency and fast sample and rendering rates, and a balance between visual and musical quality [Fels et al. 1997; Fels and Mase 1999]. Furthermore, having a direct relation between action and result and a reflection of self in the stimulus (hand shape seen in obstructions) is an innovation.

## 5. References

- FELS, S., AND MASE, K. 1999. Iamascope: A Graphical Musical Instrument. *Computers and Graphics*, 2(23), 277-286.
- FELS, S., NISHIMOTO, K., AND MASE, K. 1997. MusiKalScope: A Graphical Musical Instrument. *Proceedings of IEEE International Conference on Multimedia Computing and Systems*, 55-62.



**Figure 2:** Whole-hand input from the MTC Express creates obstructions (blue circles) that affect the flow of the particles.

# Foamy Creatures: Digital Domain wrangles whitewater for “Lord of the Rings”

Markus Kurtz  
Digital Domain

Greg Duda  
Digital Domain

This sketch explains the technical and artistic techniques used by Digital Domain to produce the *Ford of Bruinen* sequence in New Line Cinema’s *Lord of the Rings*.

In 2001, the first installment of Peter Jackson's *Lord of the Rings* trilogy was released. His vision of Tolkien's epic met with a critical and box office success. During the film's production, Digital Domain was charged with completion of the challenging *Ford of Bruinen* sequence. This sequence shows Arwyn (Liv Tyler) and wounded Frodo (Elijah Wood) on horseback, pursued by a squadron of dark, ghostly riders called Ringwraiths. As the Ringwraiths negotiate a river canyon, Arwyn summons a tidal wave of raging whitewater. Unlike a real flood, this magical wave does not spill flat across the terrain seeking the lowest point, but goes directly for the Ringwraiths in a compact, dynamic mass. The churning, frothy watery mass cascades as its front turn into a herd of horses representing the avenging spirit of the spell. The whitewater stampede rolls over the Ringwraiths causing their demise. Afterwards, the magical wave subsides in a natural fashion.

To bring this concept to life, Digital Domain had to create the following:

- a realistic flash flood moving alongside a directed path
- articulated, whitewater horses to be incorporated into the wave front
- CG stunts of the Ringwraiths as they meet their sorry fate.

The realistic look and behavior of the rapids-like whitewater was achieved with an in-house fluid dynamics simulator. The voxelized geometry of the riverbed served as the simulation environment. The simulation data was subsequently post-processed in Side Effects' Houdini in order to control the direction and to add large scale splashes at predefined dramatic points. The simulations were rendered with proprietary volumetric software. The software took into account the velocity and orientation of the

simulation to maintain a uniform look. Noise shifts and rotations were introduced to add localized dynamics. To add finer detail to the simulation, a separate, non-volumetric technique was devised. It relied on geometric primitives, such as cards and spheres, which were driven by the simulation and rendered with a shader representing foamy detail.

The creation of the whitewater horses took a twofold approach. For the chest and neck, volumetric clouds were applied to animated skeletal joints animated in Maya. For the head region, volumetric Booleans were used to carve out an accurate representation of the head's form. To achieve the effect of long, flowing manes, time lag techniques were employed which inherited and manipulated volume information between frames. The facial detail was touched up with shaded geometry layers and enhanced with independent particle systems.

At the beginning of the sequence, the doomed Ringwraiths were assembled from individual live action elements. For the falling and tumbling shots, Digital Domain created CG "stuntmen". They were animated in Maya and rendered with Renderman shaders to match Weta FX's look. Z-depth compositing allowed the Ringwraith squadron to be visually integrated with CG whitewater.

The look of the tidal wave was finessed with supplementary touches. Live action elements shot at Niagara Falls were mixed in to add local, dynamic splashes. Thin, hovering mist was emitted from the simulation's iso-surface and rendered as a separate volumetric element. The realism was further enhanced with small scale, specular detail, and hand-placed particle systems for obstacle interaction.

Judging by positive reviews, including Roger Ebert's singling out of this sequence in his praise of the movie, Digital Domain's met the challenge with success.



Figure 1. (a) 3d whitewater horse and simulation elements (b) 2d water elements (c) final composited image

# Fracture generation on polygonal meshes using Voronoi polygons

Saty Raghavachary  
DreamWorks Feature Animation

## Abstract

This sketch describes a way of generating realistic cracks and fragments to visually simulate brittle fracture on polygonal surfaces.

**Keywords:** Natural Phenomena, Visual Simulation

## 1 Introduction

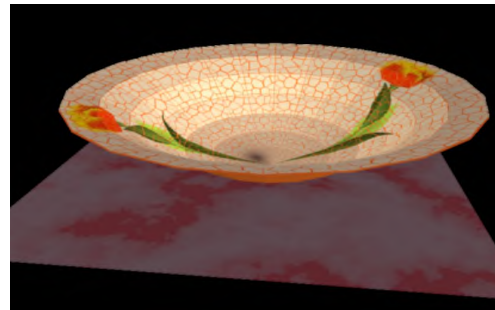
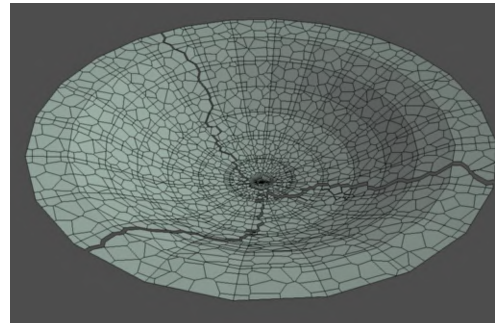
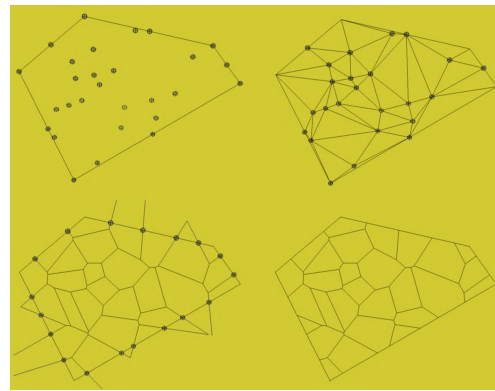
Prior computer graphical approaches to modeling inelastic deformation and brittle fracture have employed continuum approaches<sup>1</sup>, cubical lattice spring networks<sup>2</sup> and more recently, tetrahedral finite elements<sup>3</sup> and Delaunay tetrahedra coupled with distance-preserving constraints<sup>4</sup>. The approach outlined in this sketch focuses on the generation of realistic cracks and fragments through synthesis of Voronoi polygons on each face of a given polymesh and subsequent propagation of cracks through the mesh by the separation of adjacent Voronoi polygons along their shared edges<sup>5</sup>.

## 2 Cracks, fracture, fragment generation

The first step is to synthesize a Voronoi network on the polygons of the input mesh. This involves seeding the interior and edges of each polygon with additional seed points, triangulating the resulting point set (expanded to include the polygon vertices as well), constructing the Voronoi tessellation from the triangulation (its geometric 'dual'), and most importantly, clipping the Voronoi network against the edges of the parent polygon. The result is a collection of smaller polygons which tile the original polygon. The first figure on the right illustrates this step. When such a synthesis is carried out independently on each polygon of the original mesh, we obtain a 'piecewise' Voronoi network over the mesh. Note that a Voronoi network contains convex polygons where exactly three edges radiate from each vertex. This property is what is behind the visually realistic cracks and fragments we are about to generate.

The next step involves using the Voronoi polygons to cause the parent surface to crack/fracture. The idea is to advance cracks along

the edges of the Voronoi polygon networks. This can be done in a variety of ways. For instance, Voronoi vertices could nucleate cracks which would traverse the parent surface until they terminate at a free edge or at an existing crack (see the second figure below). Alternately, a small number of spatially separated Voronoi polygons could serve as seeds for region-growing which would proceed outwards from these seeds, generating fragments. Holes could be created in the surface by removing some of these regions. Alternately the surface could be shattered into very small fragments by essentially turning each Voronoi polygon into a shard. Also, a network resembling ceramic 'glaze' cracks can be obtained by slightly shrinking each Voronoi polygon about its centroid (third figure below). These approaches could be either procedurally driven or be under animator control. As a final note, each of the generated fragments could be used to create Voronoi networks on a finer scale which in turn would be the source of smaller crack networks and fragments.



<sup>1</sup>D. Terzopoulos, K. Fleischer: SIGGRAPH 1988 Proceedings, p. 269-278

<sup>2</sup>A. Norton et. al.: The Visual Computer, Vol.7, 1991, p. 210-217

<sup>3</sup>J. O'Brien, J. Hodgins: SIGGRAPH 1999 Proceedings, p. 137-146

<sup>4</sup>J. Smith, A. Witkin, D. Baraff: Graphics Interface 2000, p. 27-34

<sup>5</sup>Saty Raghavachary, M.S. Thesis (CIS Department), The Ohio State University, 1992



# Generating Feather Coats Using Bezier Curves

L. Streit\* and W. Heidrich

Imager Computer Graphics Laboratory, University of British Columbia

## 1 Introduction

Feathers, like hair and fur, dramatically alter the appearance of a surface. However, unlike hair and fur, feathers have a wide range of colours and patterns and a well-defined branching structure. The variety of individual feather structures and patterns contribute to the surface appearance. Thus, it is important to model the complete range of feather types found in a feather coat. Since a feather coat can consist of thousands of structurally unique feathers, it is desirable to automatically generate most of the feathers, while maintaining intuitive control over the coat design and the creation of a wide variety of feather types. Previous work on feathers [Dai et al. 1995] does not address the intuitive generation of such a coat.

## 2 Feather Structure

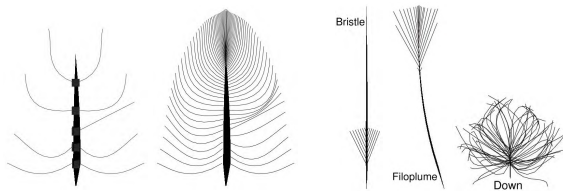


Figure 1: Left: keybarbs; Right: feather structure.

Figure 2: Variety of created feather types.

We need to automatically generate feathers that are individually unique, but that exist in a coat comprised of a certain range of feather types; therefore, the parameterization used to model the feather structure must allow for smooth interpolation between types and ages of feathers. The parameterization we propose is based on Bezier curves. Both the main shaft of the feather and the branching sub-structures (*barbs*) are defined by Bezier curves. A small set of parameters based on physical sub-structures of feathers are used. These include the barb spacing, the barb angle at the base and tip of the shaft, the length of the barbs and the length and width of the shaft. Additionally, a set of shape-defining **keybarbs** are specified. Barbs between these keybarbs are generated by interpolating the control points of the keybarbs (see Figure 1). In the absence of keybarbs, a default barb with random control vertices is used to simulate the downy portions of feathers. The generated feather types are shown in Figures 2, 4, and 6.

### 2.1 Interpolation



Figure 3: Interpolation simulating feather growth.

Figure 4: Simulated peacock feather.

In a similar manner, interpolating between keybarbs on different feathers generates a range of feather types. The number of keybarbs may be temporarily inflated to interpolate between feathers

with varying numbers of keybarbs. Using the feathers defined with keybarbs and the interpolation constants, a feather with a new set of keybarbs is generated. Intermediate barbs are generated as above.

Finally, interpolation of all the parameters, in sequence, simulates feather growth. First the shaft length/width and barb length are interpolated to their final values. This simulates the growth of an encased *pin feather*. The opening of the feather is simulated by interpolating the keybarb control vertices, from a position along the shaft to their final positions, while interpolating the shaft control vertices to their final position. Figure 3 shows an example of simulated growth.

### 2.2 Texturing and Rendering

The shaft is modelled as a generalized cylinder and the barbs are modelled using two triangle strips on either side of the Bezier curve. The barbs are then alpha-blended with specific order length-wise (from base to tip) and width-wise (from the centre to the edges) to simulate the transparency of feathers.

## 3 Feather Coats

Finally, using a small set of defined feathers (*key feathers*) positioned at particular locations on the model, bi-linear interpolation is used to generate uniquely structured feathers at all intermediate locations (see Figure 5).

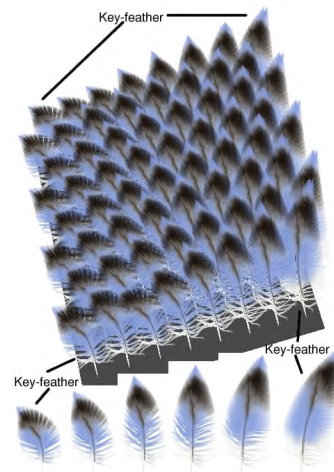


Figure 5: Feather coat example using two feather types positioned at the four corners.



Figure 6: Hawk feather; generated version on left, scanned on right.

## References

- DAI, W.-K., SHIH, Z.-C., AND CHANG, R.-C. 1995. Synthesizing feather textures in Galliformes. *Computer Graphics Forum* 14, 3 (Aug.), 407–420.
- IONES, A., KRUPKIN, A., VOLODARSKY, S., AND ZHUKOV, S. 2000. Fur and hair: practical modeling and rendering techniques. *The Proceedings of the IEEE International Conference on Information Visualization (IV'00)*.
- LUCAS, A. M., AND STETTENHEIM, P. R. 1972. *Avian Anatomy Integument*. U.S. Government Printing Office, Washington D.C.

\*e-mail: Streit@cs.ubc.ca



# “GEO-COSMOS”: World’s First Spherical Display

**CONTACT**  
Tamotsu Machida  
ViNO azul, Inc.  
mach@vino.co.jp

**CONTRIBUTORS**  
Tsuyoshi “Go” Hotta,  
DENTSU Inc.,  
DENTSU TEC, Inc.,

National Museum of Emerging Science and Innovation

## Introduction

Display technology has been advancing every year. However, the topologically speaking, almost all of these use the same "Planar" surface for their displays. The topology of our display, to be introduced here, is completely different, in that we use "Spherical" surface. GEO-COSMOS, the world's first full color spherical display, is truly remarkable. It is also the main exhibit of the National Museum of Emerging Science and Innovation, opened in Tokyo on July 10, 2001. As shown in Figure 1 and 2, the images/movies on this display are intended for the viewers from all directions.

## 1 Technical Description of GEO-COSMOS Spec

This spherical GEO-COSMOS is 6.5m in diameter, and has a total of 3715 LED (light-emitting diode) panels attached on the surface. Unfortunately we could not attach 77 panels, because the cable wiring needed the space for attaching this sphere to the ceiling. 256 SMD (Surface Mounted Devices) Type LED dots (16 dots x 16 dots) are embedded in 10mm pitch on each panel, making the entire screen consist of approximately 950,000 dots. Each LED has all of RGB components in a single dot and 10bits as the color depth. SMD type LED has many advantages over other type of LEDs. SMD type offers shorter distances for color mixing, less blurring of colors in relation to the movement of viewers, and considerable improvement in decline of luminance. Hence, the viewers can see this display as naturally as a flat LED screen.

## 2 Displayed Images

Designed to suit today's science museums, the contents are selected to show various satellite data images of the earth. At this stage, we show four kinds of data images, provided by University of Wisconsin and NASA.

1. "Earth Today (cloud cover of last 30 days)": There are four geostationary weather satellites that surround the earth. The visible and IR scanned images are sent from these satellites, and University of Wisconsin makes them into a composite in a cylindrical projection and sends them to the museum in Tokyo via the Internet every 3 hours.

2. "Carbon monoxide concentration (2000/3/1 - 2000/12/30)": Provided by NASA

3. "Earth today (showing the night and day boundary)": Provided by University of Wisconsin and NASA

4. "The museum logo animation": We created this spherical 90-seconds-logo animation, interpreting and expressing the museum logo artistically and conceptually. This is also the first real spherical CGI animation in the world (Figure3-5.)

In addition, the contents include the lunar phases, like waxing and waning of the moon, which is shown during the moon-watching events in traditional Japanese calendar. Currently, all contents are pre-rendered and stored in HD DDR (High Definition Digital Disk Recorder). They are displayed on GEO-COSMOS on a preset schedule.

## 3 Image Processing

As long as the data is available to be mapped into a spherical shape (that is called cylindrical projection), any image data can be dis-

played. The data is first calculated in GNU/Linux servers to match the LED chip positioning and then is stored in HD DDR. Making a simple image composite, displaying text, and rotating the image can also be done in this image processing. This proprietary software is written in C++. The stored video format is standard SMPTE 274M 1080/30p. Fiber optic cables connect HD DDR and LED processor, inside of GEO-COSMOS.

## 4 Future Work And Conclusion

Currently we plan to add more contents and to develop capabilities for interactive/real-time display. It is logistically impossible to bring GEO-COSMOS to the SIGGRAPH conference, so we hope you will come to the museum and to experience GEO-COSMOS.

## Acknowledgements

We would like to acknowledge the NASA Goddard Space Flight Center, the Space Science and Engineering Center University of Wisconsin- Madison, the American Museum of Natural History, the Rose Center for Earth and Space American Museum of Natural History.



Figure 1. Viewing from 5th floor.



Figure 2. Viewing from 1st floor.



Figure 3. Logo animation(1).

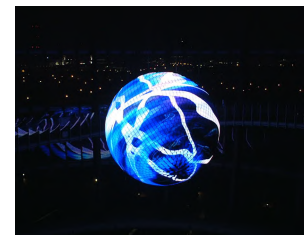


Figure 4. Logo animation(2).



Figure 5. Logo animation(3).

# Hardware-accelerated Texture and Edge Antialiasing using FIR Filters

Koen Meinds  
Philips Research  
koen.meinds@philips.com

Bart Barenbrug  
Philips Research  
bart.barenbrug@philips.com

Frans Peters  
Philips Research  
frans.peters@philips.com

## 1 Introduction

Texture mapping is a standard feature of 3D graphics systems. To avoid aliasing artifacts, proper filtering is mandatory. We have developed a novel algorithm for texture mapping and filtering that is suited for hardware implementation. To eliminate texture aliasing artifacts, our algorithm uses higher order FIR filters that are known from digital signal processing. Compared to current state-of-the-art “anisotropic level 8 filtering”, that can be found in commercially available advanced graphics accelerators, our method produces higher image quality at equal costs.

Apart from aliasing in transformed textures, aliasing can also occur at the edges of polygons (staircase jaggies). We have combined our filtering algorithms with an A-buffer like approach to prevent edge aliasing. Compared to current state-of-the-art supersampling or multisampling, our method produces higher image quality than 4 x 4 supersampling at costs similar to 2 x 2 supersampling.

## 2 Approach

Unconventionally, we rasterize the polygons in texture space and traverse the texels inside the polygon boundaries. The contribution of each input sample is “splatted” to an area of output samples (pixels): eg. the texel at the red arrow in Figure 1 is splatted to the bold dots.

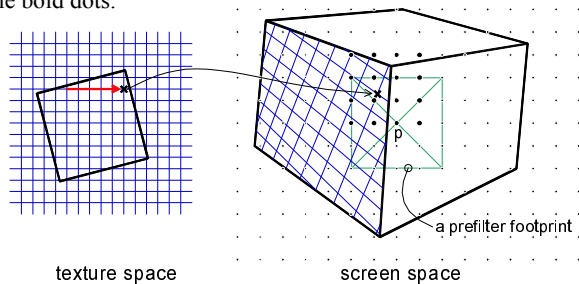


Figure 1: Forward texture mapping

We use the following filter formula:

$$I_p = \sum_t h(X_p - X_t) f_t I_t \quad (1)$$

It gives the color  $I_p$  of a pixel  $p$  with coordinate  $X_p$  depending on the colors  $I_t$  of texels  $t$  with mapped coordinate  $X_t$  that are within the pre-filter footprint of  $p$ ;  $f_t$  is the so-called local scaling factor. Conventional mipmaps are used to select convenient resolutions for sampling in texture space.

To obtain optimal results, the pre-filter function  $h$  needs to be chosen carefully; there are trade-offs between sharpness, blurriness, aliasing and ringing. We selected a function of four pixel spacings with negative lobes.

The input driven filter technique, that “splats” the texel colors to the pixels, is not only suited for texture antialiasing but can easily be extended to edge antialiasing, as will be explained next. All pixels whose pre-filter footprint covers the mapped texel

under consideration must receive a contribution from the texel. Hence, also pixels that fall outside the polygon that is being rasterised, may receive contributions (eg. the pixel with label  $p$  in Figure 1). To determine the final color of pixels around edges, partial colors of all contributing polygons are required. We therefore store the generated partial colors in a *fragment buffer*; this buffer replaces the traditional Z buffer. Each *pixel fragment* contains the partial color together with other data. The fragment buffer is similar to the A-buffer, but we do not use subpixel masks.

## 3 Results and future work

Our results exhibit better texture antialiasing compared to anisotropic filtering that can be found in current advanced graphics chips. Our edge antialiasing results show absence of aliasing on most edges (edges that result from polygon intersection are not yet dealt with). Our results are most appreciated when viewing animated sequences and comparing them with the results from traditionally hardware rendered images. Figure 2 shows a shot from such a sequence, comparing our (right) with the traditional non-supersampled (left) method.

This sketch reports on on-going work. Results thus far have been obtained by one pass texture mapping, i.e. the filter footprints are two-dimensional. There are, however, many advantages in using two-pass texturing: first a horizontal pass followed by a vertical pass. We are investigating the two-pass algorithms. Moreover, we are analysing the various trade-offs that exist when realising the algorithms in hardware.

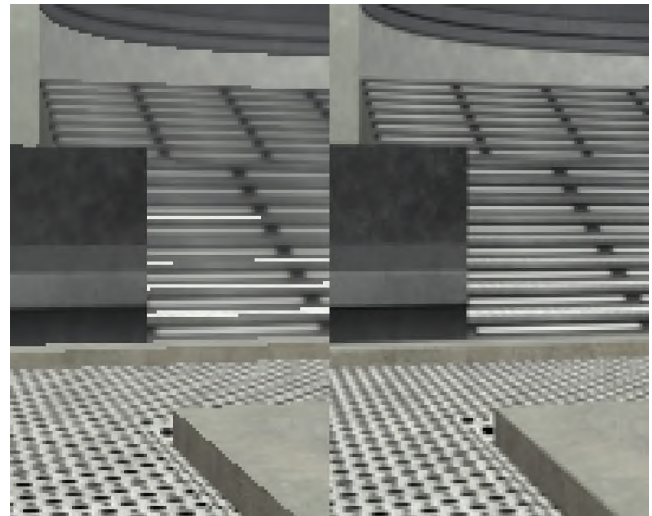


Figure 2

**Acknowledgement:** We thank the owners for their permission to use their models and textures in our demonstration sequences; see: <http://bettenberg.home.mindspring.com/teqtrny3.html> and <http://www.planetquake.com/hfx/>.

# Harmonic 3D Shape Matching

Michael Kazhdan, and Thomas Funkhouser \*  
Princeton University

## 1 Introduction

With the advent of the world wide web, the number of available 3D models has increased substantially and the challenge has changed from “How do we generate 3D models?” to “How do we find them?” In this sketch we describe a new 3D model matching and indexing algorithm that uses spherical harmonics to compute discriminating similarity measures without requiring repair of model degeneracies or alignment of orientations. It provides 46–245% better performance than related shape matching methods during precision-recall experiments, and it is fast enough to return query results from a repository of 20,000 models in under half a second.

## 2 Model Representation and Matching

The main challenge in designing a model matching algorithm is to find a computational representation of shape for which an index can be built and geometric matching can be performed efficiently. Generally speaking, the following properties are desirable for a shape representation. It should be: (1) quick to compute, (2) concise to store, (3) easy to index, (4) invariant under similarity transforms, and (5) independent of 3D object representation, tessellation, genus, or topology.

Previous work in defining shape representations for matching can be classified into two broad categories. The first class of representations use methods such as PCA to align the model into a canonical coordinate frame and then define the shape representation with respect to this orientation. Such methods include, among others, Moments [Elad et al. 2001] and Extended Gaussian Images [Horn 1984]. The second class of methods define representations that are invariant under rotation and include methods such as Shape Histograms [Ankerst et al. 1999] and Shape Distributions [Osada et al. 2001]. In our experiments we have found that PCA based methods are unstable as a result of multiplicity of eigenvalues and sensitivity to outliers. The shape representation that we have designed is both rotation invariant and discriminating, characterizing a shape in terms of the clustering of mass on different concentric spheres.

The steps for computing the harmonic shape representation are outlined in Figure 1: (1) Given a model, we rasterize its polygons into a  $64 \times 64 \times 64$  voxel grid, (assigning a voxel a value of 1 if it was within one voxel of a point on the boundary, and a value of 0 otherwise). The model is aligned so that its center of mass is at the center of the grid, and so that its bounding sphere has radius 32. (2) Treating it as a function defined in three-space, we decompose the voxel grid into 32 spherical functions by restricting the voxel grid to spheres with radii 1 through 32. (3) We decompose each of these functions as a sum of its first 16 harmonic components, analogous to a Fourier decomposition into different frequencies. (4) Using the fact that rotations do not change the norm of the harmonic components, we define the signature of each spherical function as a list of these 16 norms. (5) Finally, we combine these different signatures to obtain a  $32 \times 16$  signature for the 3D model. The resultant rotation invariant signature is a two-dimensional grid where the value of the  $(i, j)$ -th index is equal to the norm of the  $j$ -th order component of the spherical function on the sphere of radius  $i$ .

To compare two harmonic representations, we simply compute the Euclidean distance between them. Thus, finding the  $K$  closest models to a query is equivalent to solving the nearest-neighbor problem

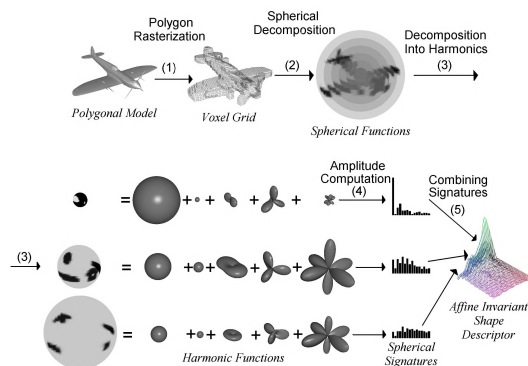
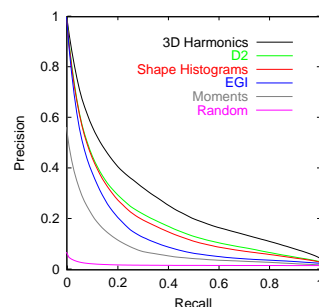


Figure 1: Computing the Harmonic Shape Representation

in a  $32 \times 16$  dimensional space. Although this problem is known to be hard in the worst case, we can build a search algorithm that works efficiently in practice by taking advantage of the multiresolution nature of the harmonic decomposition to guide a dimension reduction indexing method.

## 3 Results

We compared the shape classification performance of our method to five existing methods, using a test database of 1890 models provided by Viewpoint. The models were clustered into 85 classes based on functional similarities. The smallest class had 5 models, the largest 143, and 610 models did not fit into any meaningful class. The other methods tested were: Moments, Extended Gaussian Images, Shape Histograms, and D2 Shape Distributions. The figure on the right shows the precision vs. recall graphs for each method. Note that our method has precision values on average 42% higher than D2, 60% higher than Shape Histograms, 126% higher than EGIs, and 245% higher than moments.



## 4 Conclusion

In summary, this sketch investigates a new method for model matching and indexing. The main research contribution is a new shape representation that allows for efficient, robust, and discriminating querying of models in large databases.

## References

- ANKERST, M., KASTENMÜLLER, G., KRIEGEL, H.-P., AND SEIDL, T. 1999. 3d shape histograms for similarity search and classification in spatial databases. In *Proc. SSD*, 207–226.
- ELAD, M., TAL, A., AND AR, S. 2001. Content based retrieval of vrmf objects - an iterative and interactive approach. *EG Multimedia* (September), 97–108.
- HORN, B. 1984. Extended gaussian images. *Proc. of the IEEE* 72, 12 (December), 1671–1686.
- OSADA, R., FUNKHOUSER, T., CHAZELLE, B., AND DOBKIN, D. 2001. Matching 3d models with shape distributions. *Shape Modeling International* (May).

\*e-mail: mkazhdan, funk@cs.princeton.edu



# Hiding Spaces: a CAVE of Elusive Immateriality

Cynthia Beth Rubin  
Rhode Island School of Design  
cbrubin@risd.edu

Daniel F. Keefe  
Brown University  
dfk@cs.brown.edu

## 1. Introduction

*Hiding Spaces* is an immersive VR Cave artwork which pushes past the limitations of physical media by exploring the new ambiguities that can delight the viewer in the virtual world. By using innovative tools developed especially for creative work within the Cave environment, in combination with more established digital methods and artistic practice, the authors collaborated to produce a work which transgresses the usual borders of 2D and 3D, including those that are common even in VR environments.

## 2. Description

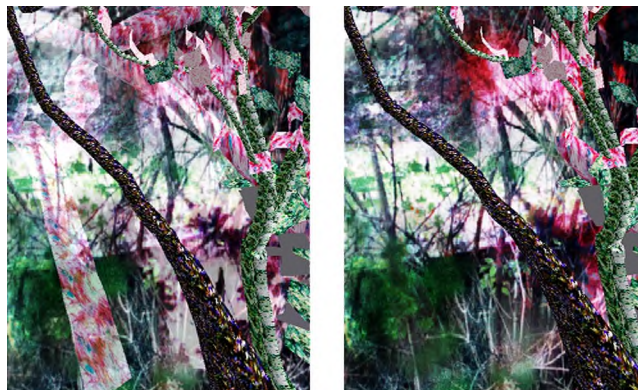
Engaging visual art depends on a rhythmic interplay of texture, color, and form to captivate the viewer and to create the sensation of movement. Many artists, particularly those working in static 2D media, employ these elements to create ambiguous spatial tensions. Artists love ambiguity, and 2D artists especially are enamored of spaces that are never defined but rather remain elusive, because spatial illusion is what makes traditional 2D imagery come alive.

In the Cave environment we have the opportunity to further the power of elusive space by incorporating the play of “real” space against the illusion of ambiguous space. In this work, the 2D spaces breathe not only because of the illusion of complex spatial relationships, but through the actual non-static nature of the projections. Rather than deny the limitations of the physical walls of the Cave, *Hiding Spaces* utilizes them as an integral part of the work.

*Hiding Spaces* is comprised of 2D painterly images in combination with virtual 3D objects, using software that runs in a 4-wall immersive virtual reality system (Cave). To experience the work, viewers wear stereo shutter glasses with a positional tracker. As the viewer moves around the space, the system detects each new position, thus triggering corresponding shifts in the imagery. While the 3D forms remain anchored at the walls, the 2D images, projected on the walls and floor, slowly shift by means of a soft cross-fade.

The entire human-computer-interaction is based on the familiar and natural activities of walking and looking. The viewer is drawn to a 3D object or a spot of 2D imagery, but in moving towards it, finds that this very action prompts a change. When something else catches the viewer’s attention, the cycle repeats. The user is constantly compelled to explore, yet doesn’t have to learn any complex interaction modalities. By simply walking around and exploring naturally, the viewer quickly learns that it is engagement and curiosity that cause the imagery to shift, and this in turn encourages more interaction with the system.

*Hiding Spaces* evolved from photographic sources manipulated in Adobe Photoshop and Corel Painter, first used in an earlier 2D printed digital image. As we discovered that projecting shifting images on walls was too jarring without slowly blending transitions, we developed a method for creating convincing digital color and texture in varying overlays. The proprietary software CavePainting<sup>1</sup> allowed us to create forms within the Cave itself in response to the imagery on the wall. As a result of our experiments with the juxtaposition of 2D and 3D elements in the Cave, we eventually developed textures for the 3D CavePainted forms out of image fragments of the projected 2D images.



The images shown here depict views of 3D forms against one wall of the Cave. In the example above, we see a cluster of pink forms gathered at various angles in 3D space. In the image on the left, a tree-like 3D form with similar coloring is just fading away but still present. The pink forms relate visually to this 3D form, as well as to the rich pink-purple texture on the wall. As the viewer moves, the situation shifts slowly, so that by the time we get to the instance on the right, the cluster of 3D pink forms appears to be less intense, even though in fact the forms are still there. The clearer background, with an intense red spot, contributes to the changing perception, as does the disappearance of the pink tree form. In the examples below, we see more clearly how the changing background frames the reading of the 3D forms.



## 3. Conclusion:

Engaging visual art, in the Cave as well as in static media, depends on a rhythmic interplay of texture, color, and form to captivate the viewer and to create the sensation of movement. But the Cave environment introduces new notions of materiality and non-linear space. Thus how we develop and apply innovative tools for creating within this environment, as well as how we use familiar tools, can be instrumental in developing a new aesthetic. Ultimately, we envision that the lessons learned through artistic work in the Cave will also be applied to problems of visual representation in scientific domains.

Thanks to the Brown University Graphics Group, Scientific Visualization Group, and the Technology Center for Advanced Scientific Computing and Visualization, and to David Laidlaw and Sean Solley for their help in this project. This work was partially supported by NSF (CCR-0086065, EIA-9871440).

<sup>1</sup> Keefe, D. et al. 2001 CavePainting: A Fully Immersive 3D Artistic Medium and Interactive Experience, In *Proceedings of ACM Symposium on Interactive 3D Graphics 2001*, J. Hughes and C. Séquin Ed., ACM Press / ACM SIGGRAPH, New York, ACM, 85-93



# High-speed Conversion of Floating Point Images to 8-bit

Bill Spitzak, Digital Domain \*

## Introduction

Most rendering software today destroys their accurate lighting and shading calculations by doing an inaccurate linear conversion to a screen image. This sketch presents a technique<sup>1</sup> that quickly converts floating point data to a screen image while preserving the correct brightness levels and original detail.

## Screen Gamma

A common misconception of how monitors work is the amount of light emitted by the screen is in linear proportion to the number placed into the image buffer, and that multiplying floating point pixel colors by 255 will produce the screen image. However a power function is the correct mapping and has been standardized[Int n. d.]. The proper way to convert a floating point brightness is to use the inverse of sRGB and multiply that by 255:

$$to\_sRGB(x) = 255 \times \begin{cases} 12.92x & \text{if } x \leq .04045/12.92 \\ 1.055x^{1/2.4} - .055 & \text{if } x > .04045/12.92 \end{cases}$$

It may also be desirable to present high dynamic range image data by compressing the high end. There are many ways to do this. One possibility is:

$$to\_sRGB\_compressed(x) = \begin{cases} to\_sRGB(x) & \text{if } x \leq .5 \\ to\_sRGB(1 - 1/4x) & \text{if } x > .5 \end{cases}$$

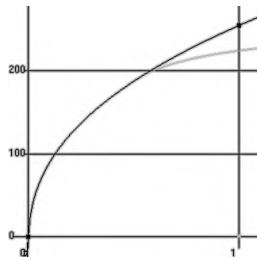


Figure 1:  $to\_sRGB$  (black) and  $to\_sRGB\_compressed$  (gray)

Since the domain of  $to\_sRGB\_compressed()$  is from zero to infinity we would like our conversion to be able to handle this.

## Look-up Table

A look-up table indexed by the top 16 bits of the floating point numbers is used to avoid the need for conditional expressions or calls to `pow()`. This table, however, is quite large as there are 65536 entries. If this is a concern the table can be made smaller by adding range checks to the algorithm. Limiting the table to  $.04045/12.92 \leq x \leq 1$  (and linearly interpolating outside that range) would reduce the number of entries to 1065.

To fill the table a floating point number is constructed with the correct high 16 bits and the lower bits set to 0x8000, which is halfway through the range that maps to that table entry. This requires care on processors that trap NaN or otherwise fail on some bit patterns in floating point. Some compilers may also have trouble

\*spitzak@d2.com

<sup>1</sup>GPL code is available at <http://www.cinenet.net/spitzak/conversion>

loading arbitrary bit patterns into floating point registers. The floating point value is then converted to a byte, rounded and clamped, and stored in the table.

How accurate is this table? The distance between output values of the table is  $\Delta y(x) = 2^{\lfloor \ln x / \ln 2 \rfloor - 7} d(to\_sRGB)/dx$ . The maximum  $\Delta y(x)$  when  $0 \leq x < 1$  is at .5 and is  $\approx .656$ . Thus every possible byte value is represented, and if values are calculated at the center of the table entries the maximum error is less than 1/3 of an output value. At common image levels the error is smaller, at 18% gray the maximum error is .15.

## Error Diffusion

Floating point data can represent fine gradations in color that are lost if the values are simply rounded to the nearest byte value. An error diffusion algorithm is used to solve this. Most algorithms are designed to reduce the image to very few levels and involve the weighted distribution of "error" to four or more neighboring pixels[Floyd and Steinberg 1975]. With 256 output levels, it is sufficient to transfer all the error to the next pixel. This allows the conversion of each scan line to be independent and is of course very simple.

The error information is stored by making the table contain 16-bit words equal to  $256 \times \lfloor to\_sRGB(x), 255 \rfloor$ . The top 8 bits are then the byte value lower than the desired result and the lower 8 bits are the error offset. Starting the error at 0x80 results in rounding to the nearest byte.

This simple dithering would produce vertical lines in solid shades and horizontal ramps. This was solved by dithering in both directions away from a randomly chosen starting location. Solid areas of black or white must not set the error back to zero or vertical lines will appear in the regions after those.

## Preserving 8-bit Data

It is desirable to have 8-bit files converted to floating point and back to emerge unchanged. If this was not done then repeatedly reading and writing an 8-bit image would gradually deteriorate it. To enforce this the table is modified: each 8-bit pixel value is converted to floating point and the resulting table location is replaced with the exact integer. Since the error is zero, no dithering happens when writing the data back out. This reduces the accuracy of the table by spacing the samples somewhat irregularly. However no entries are out of order and only 1.6% of the table entries are changed. The result is that only a small amount of noise is added to the data.

## Conclusions

This algorithm has proven to be fast enough for real-time preview of compositing results. It has also proven that trivial calculations of blurs and exposure changes in linear floating point produce beautiful and photographically correct results when converted this way. This technique promises to allow future software to produce physically accurate pictures quickly and easily.

## References

- FLOYD, R., AND STEINBERG, L. 1975. An adaptive algorithm for spatial gray scale. *Society for Information Display 1975 Symposium Digest of Technical Papers*, 36.
- INTERNATIONAL ELECTROTECHNICAL COMMISSION. *The sRGB Specification*, IEC 61966-2-1 ed.

# Hover: Conveying Remote Presence

**Dan Maynes-Aminzade**  
MIT Media Lab  
One Cambridge Centre  
Cambridge, MA 02139  
monzy@media.mit.edu

**Beng-Kiang Tan**  
Harvard Graduate School of Design  
48 Quincy Street  
Cambridge, MA 02138  
btan@gsd.harvard.edu

**Ken Goulding**  
MIT Department of Urban Studies & Planning  
77 Massachusetts Ave  
Cambridge, MA 02139  
goulding@mit.edu

**Catherine Vaucelle**  
MIT Media Lab  
20 Ames Street  
Cambridge, MA 02139  
cati@media.mit.edu

## Abstract

This sketch presents *Hover*, a device that enhances remote telecommunication by providing a sense of the activity and presence of remote users. The motion of a remote persona is manifested as the playful movements of a ball floating in midair. *Hover* is both a communication medium and an aesthetic object.

## 1 Introduction

Face-to-face dialogue can be more engaging than telephone conversation because of the added elements of gesture, touch, and body language. Videoconferencing attempts to provide these missing elements, but does so at the cost of high bandwidth, expensive equipment, and heightened demands on the attention of the user. *Hover* provides a low-cost, low-bandwidth, less distracting solution that enhances the experience of telephone conversations with family and friends. It provides visual awareness of remote persons in the form of an abstract physical representation with several affordances: a real-time indication of the level of physical activity of the remote person; the ability to personalize the representations of the remote persons in a way that makes sense to the user; and the ability to grasp and interact with a “surrogate” representing the remote person. *Hover* is not intended to convey the meaning of gestures in a conversation, but rather to convey a sense of presence of the remote person in a captivating and poetic fashion. We were inspired by other works that used physical objects to communicate, to show activity or presence, or to support intimacy [Strong and Gaver 1996; Ishii et.al. 2001; Ishii and Ulmer 1997].

## 2 Description

*Hover* uses technology that can easily operate in a home or office environment. It is connected to a computer and telephone or an Internet phone. The user places multicolored balls representing persons with whom she frequently communicates (e.g. family and friends) on a stand on the *Hover* platform. She can personalize the identity of the balls by painting different colors and patterns on them. In a scenario where Peter calls Jane (refer to <http://www.media.mit.edu/~monzy/hover.swf> for an animation of this process), Peter first takes a ball representing Jane and puts it on the top of the ramp. The ball is sensed, then rolls down and stays at the foot of the ramp. Jane’s number is then dialed automatically. On Jane’s end, the ball that represents Peter floats when the call comes in. If Jane wants to pick up the call, she grabs the ball and puts it on the ramp (Figure 1). The ball rolls across a sensor on the ramp, sending a signal to Peter’s end to indicate that Jane has picked up the call. At the foot of the ramp, an air stream levitates the ball (Figure 2). As Peter speaks, Jane sees the ball floating up and down in correspondence with Peter’s movements (Figure 3). This will convey to Jane a sense of Peter’s presence and level of activity. If Peter is inactive, the ball will hover at a fixed height. When the telephone conversation ends and both parties hang up the telephone, the ball stops floating and the user places it in its original position.

Our current implementation uses a camera with vision tracking, attached to a computer. The computer analyzes the video stream and detects the level of motion. It sends this information to the remote computer, which relays the motion data to the remote

*Hover* device. The *Hover* unit reads the data and sends the appropriate electrical pulses to the fan and servo motor to control the position and the height of the ball. If the remote person gestures wildly, the ball floats rapidly up and down.

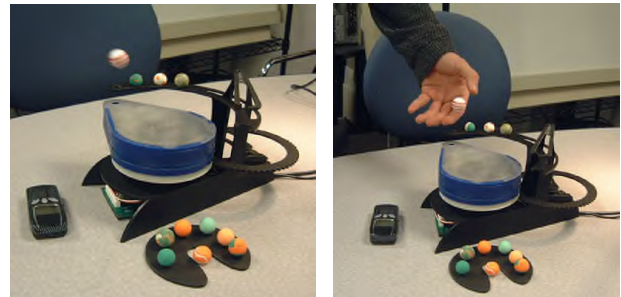


Fig. 1. At Jane’s end, the ball floats when a call comes in. She grabs the ball to answer the call.



Fig. 2. The ball rolls down the ramp and an air stream levitates it.



Fig. 3. In conversation: the ball floats up and down in correspondence to the remote person’s movement.

## Acknowledgments

We thank Professor Hiroshi Ishii of MIT Media Lab for his support and the students of his Tangible Interfaces course for their feedback.

## References

- STRONG, R., AND GAVAR, W.W. 1996. Feather, scent and shaker: Supporting simple intimacy. In *Proceedings of CSCW 1996*.
- ISHII, H., REN, S. AND FREI, P. 2001. Pinwheels: Visualizing Information Flow in an Architectural Space. In *Extended Abstracts of CHI 2001*, ACM Press, 111-112.
- ISHII, H., AND ULMER, B. 1997. Tangible Bits: Towards Seamless Interfaces between People, Bits, and Atoms. In *Proceedings CHI 1997*, ACM Press, 234-241.

# How a CSG-Based Raytracer Saves Time: Lighting and Scripting for *Ice Age*

David Esneault  
Sr. Technical Director  
Blue Sky Studios

Mitch Kopelman  
Sr. Technical Director  
Blue Sky Studios

Jodi Whitsel  
Sr. Technical Director  
Blue Sky Studios

At Blue Sky Studios, two overriding principles have guided the development of the renderer and production lighting tools: 1) the lighting model should be as physically accurate as possible, and 2) be straightforward and easy to use so that the computers take care of the technical work leaving the artists free to concentrate on the creative aspects of lighting a scene. Blue Sky's proprietary renderer, CGISudio™, has one of the most robust lighting models in the industry. The renderer's realistic approach to how light actually behaves in the real world lets the artist add details like soft shadows, reflections, and radiosity at the flip of a switch. Artists are able to achieve complex and subtle lighting with relatively simple lighting rigs, allowing them to get the image as 'right' as possible in the original render.

CGISudio™ allows the user to create many realistic light types such as ambients, disks, spotlights, and point lights with accurate dropoff rates. Global illumination solutions such as radiosity, ambient shadows (a light quality which simulates some effects of radiosity without the high cost), and lens flare give the user a powerful set of tools with which to quickly light shots. The lighting scripts used at Blue Sky, while surprisingly simple text files, still retain all the power and flexibility of the CGISudio™ scripting language.

The efficiency of the renderer can be seen in all the departments that use the software. The modeling department takes advantage of several features of the renderer. Since the software tracks directly to the surfaces of nurbs geometry, skipping the memory intensive process of polygon tessellation, less burden is placed on the modelers to create "light" geometry. The addition of subdivision surfaces to the renderer further lessens this burden. Additionally, the constructive solid geometry(CSG) base of the render engine gives the user the ability to create very complex boolean relationships between pieces of geometry and gives the user the ability to model procedurally.

In the realm of materials, the efficiency of CGISudio™ is readily apparent. With physically accurate refractions and reflections available to the user, no time is wasted rendering extra passes to generate a cheat. Secondly, the user can create material descriptions (shaders) in an interpreted language. No need to keep recompiling, once again saving time for the user. Thirdly, the lighting model of the renderer is physically accurate. Scientifically determined numbers for specularity, index of refraction, or roughness create the expected results. This means less time guessing at values for materials and more time perfecting them.

For the effects team, CGISudio™ is critical. With the ability to volumetrically render particle data through the use of a voxel-based data structure, the team has little problem generating smoke and dust. The effects department pushes the language to the limit, procedurally animating complex physical effects and managing memory for rendering extremely large data sets like the forests in *Ice Age*.

It is in the lighting process that the CGISudio™'s power really shows. The efficiency of the software allows the technical director to think like a cinematographer, not a computer scientist. The renderer requires relatively few light sources to create extremely rich imagery. In *Ice Age*, the average scene used only four or five different lights. By using concepts such as global illumination and area lights, we can achieve visually complex imagery very quickly.

Quickrender is the primary lighting pre-visualization tool at Blue Sky. It offers a front end user interface that hooks into the CGISudio™ language to allow the user to quickly visualize any given scene. Quickrender stores image data that can be recycled, so that only new information is calculated when rendering a scene. It offers near real-time feedback for basic lighting tasks, and much faster turn around than full renders for more complex lighting operations. To amplify Quickrender's speed, it is multi-threaded to run on any number of processors at once. Because it's built upon the studio language, it can also be used as a powerful development and debugging tool. Quickrender is useful for nearly all parts of the technical pipeline, allowing the user to alter scripts containing lighting, materials, modeling, and effects information. Through a combination of the power and speed of Quickrender and the straightforward nature of CGISudio™ itself, a relatively small team of technical directors were able to complete the lighting of *Ice Age* on time, on budget, and with relatively little overtime and stress that usually accompanies a production of this scale.

Clockwise from top left: Braun Shaver, *Bunny*, Quickrender Screenshot, and *Ice Age*



# How to Dress Like a Jedi: techniques for digital clothing

Ari Rapkin  
Industrial Light + Magic  
ari@ilm.com

## 1 Introduction

This sketch presents how cloth simulation is used at Industrial Light + Magic to create clothing for digital characters in Star Wars: Episode 2 and other films. Many costumes have to match complex physical costumes worn by live actors in other shots. Digital costumes have not only to look the same as their physical counterparts, but also move identically in response to their wearers' movements and environmental influences..

## 2 Multi-layer costumes

Many of our characters wear costumes which consist of multiple layers, each made of a different type of cloth with different physical characteristics. In our simulation pipeline, inner garments are simulated first, then the outer garments collide against the results of the earlier simulation. Care must be taken so that the inner garments do not reach extreme poses that the outer garment cannot enclose. For example, the sleeve of an inner shirt should not swing further from the arm than the outer robe can stretch.

## 3 “Kenny poses” and prerolls

Real-world clothing is made from flat pieces of fabric. Darts, gathers, and other devices give the garment its three-dimensional shape and affect its movement. To ensure that digital cloth exhibits the same shape and behavior, CG garments must also be constructed from segments which correspond to flat pieces of cloth. However, it's difficult for artists to sculpt CG garments in their settled, gravity-stretched poses without sculpting in some non-planar folds and puckering. To alleviate this problem, we developed a technique we call “Kenny poses” (in reference to the well-known animated character they resemble). The garment begins in a pristine state, with each section of cloth as geometrically simple as possible. We then simulate several frames of the cloth, while the character remains motionless or moves slowly into the position it will have at the start of the shot. This *preroll* allows the garment to settle, stretch and fold into a realistic position before the shot animation begins.

## 4 Multiple wind sources

In many shots, our digitally-dressed characters are standing in the wind, riding in open vehicles, and even flying or falling long distances. For these situations we need to apply simulated wind coming from various directions at changing speeds. We use multiple directional, scalable wind sources, each with its own wind speed.

## 5 Simulation controllers

Real-world costumes typically have numerous small features such as collar stiffeners, elastic insets, or weighted hems. To create these

features in a digital costume would be prohibitively complicated, but all that's really needed is to imitate the behavior the features induce. Our menagerie of simulation controllers allows animators to reproduce these effects. They include controls to stiffen a section of cloth, make it more stretchy, make it wind-resistant (we call this one “spray starch”), change its density, or make it “sticky”, so that it clings to whatever it's in contact with.

## Acknowledgements

These systems and tools were developed by ILM's software R&D simulation group, with great help and patience from the Episode 2 and Harry Potter 2 cloth-sim teams, especially James Tooley, Juan-Luis Sanchez, and Dennis Turner.vb



# ICARUS: Interactive Reconstruction from Uncalibrated Image Sequences

Simon Gibson, Jon Cook, Toby Howard and Roger Hubbard\*  
Advanced Interfaces Group,  
Department of Computer Science,  
University of Manchester, UK

## 1 Introduction

The ICARUS system is a suite of software packages, developed at the University of Manchester, that allows geometric models to be quickly and easily reconstructed from image and video sequences captured with uncalibrated digital cameras. The system combines automatic and semi-automatic camera calibration algorithms with an easy-to-use interactive model-building phase (Figure 1). Surface textures are automatically extracted from images and mapped onto the reconstructed models.

Previous algorithms for model reconstruction from images have taken two different approaches. Fully automatic algorithms (e.g. [Pollefeys et al. 1999]) use a large number of closely-spaced images to reconstruct a dense point-cloud and highly detailed triangular mesh. Approaches such as this have the benefit of ease of use, but the models they produce lack structure and contain missing regions caused by object occlusion. By way of contrast, semi-automatic approaches (e.g. [Debevec and Malik 1996; Hakim 2000]) rely on user assistance to calibrate the cameras and build scene structure, but have the benefit of being able to incrementally construct an object hierarchy. They also take advantage of user knowledge when reconstructing the environment, thereby overcoming problems caused by object occlusion. The main disadvantage of semi-automatic systems is the amount of user interaction required to identify enough common features between the images so that calibration may be accurately performed.

## 2 The ICARUS System

The approach taken by the ICARUS system falls between these automatic and semi-automatic extremes, and its robustness and ease of use comes from taking advantage of the strengths of both types of algorithm. The main source of input data for the ICARUS system is video sequences captured with a digital video camera. We combine automated feature selection and tracking with robust hierarchical structure-from-motion and novel self-calibration algorithms to automatically estimate the camera parameters for each frame of the sequence. This calibration data is then used to assist the user in interactively reconstructing a model of the environment. The new calibration algorithm has been found to be robust for a wide variety of different camera motions and sequence lengths, and has also been used for augmented video-production. Reconstruction is achieved by manipulating the position, orientation, and size of primitive objects such as polygons, boxes, cylinders etc. so that their projections into the sequence match the projections of real objects. Manipulation is aided by a variety of user-specified constraints, ranging from strictly-enforced hierarchical parent-child relationships to image-based constraints on the location of object vertices projected into each frame. Image-based constraints are evaluated using the information obtained during the calibration stage. We employ a non-linear optimization algorithm that is capable of manipulating the parameters of these objects in real-time, so that the constraints specified by the user are best satisfied.

This novel approach of combining automatic calibration of video sequences with interactive geometry reconstruction allows the user to spend more time modeling important features of the scene, rather than preparing the system for calibration. The system is not limited to video sequences, however, since we also take advantage of alternative calibration techniques for video sequences captured with tripod mounted cameras, as well as for small collections of photographs, and even single images. This gives the user maximum flexibility in terms of how the input image data can be captured, without the need to employ different reconstruction algorithms for each type of input. Example reconstructions from a variety of sources are shown in Figure 1. Typical reconstruction times for each model were less than 30 minutes. Further examples, and an evaluation copy of the software, may be downloaded from <http://aig.cs.man.ac.uk/icarus>

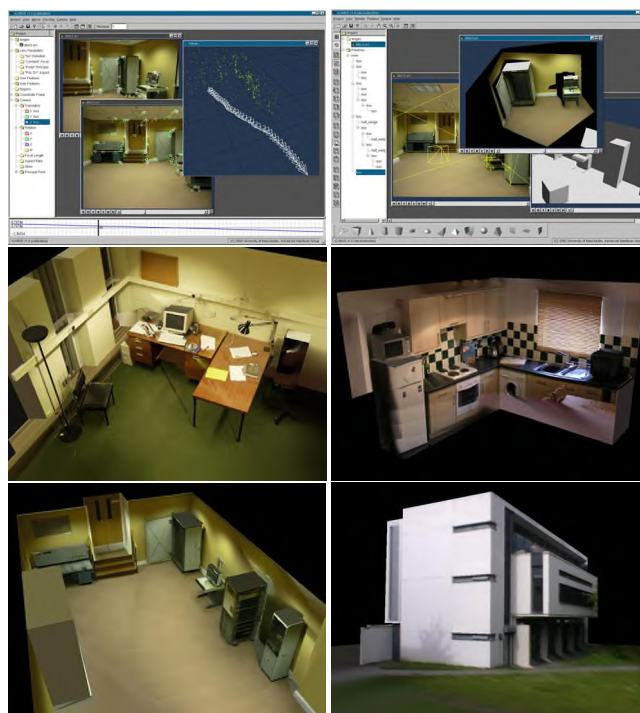


Figure 1: The ICARUS calibration and reconstruction user-interfaces (top row). Example reconstructions from a single image (middle-left), a video camera mounted on a tripod (middle-right), and from hand-held video camera footage (bottom row).

## References

- DEBEVEC, P., AND MALIK, J. 1996. Modeling and rendering architecture from photographs: A hybrid geometry and image-based approach. In *Proceedings of SIGGRAPH 1996*, 11–20.
- HAKIM, S. E. 2000. A practical approach to creating precise and detailed 3d models from single and multiple views. *Int. Archives of Photogrammetry and Remote Sensing* 33 (July), 122–129.
- POLLEFEYS, M., KOCH, R., AND GOOL, L. V. 1999. Self-calibration and metric reconstruction in spite of varying and unknown internal camera parameters. *Int. Journal of Computer Vision* 32, 1, 7–25.

\*e-mail: {sg|cook|toby|roger}@cs.man.ac.uk

# Image-based Environment Matting

Yonatan Wexler, Andrew W. Fitzgibbon, Andrew Zisserman\*  
The University of Oxford

Environment matting is a powerful technique for modelling the complex light-transport properties of real-world *optically active elements*: transparent, refractive and reflective objects. Zongker *et al* [1999] and Chuang *et al* [2000] show how environment mattes can be computed for real objects under carefully controlled laboratory conditions. However, for many objects of interest, such calibration is difficult to arrange. For example, we might wish to determine the distortion caused by filming through an ancient window where the glass has flowed; we may have access only to archive footage; or we might simply want a more convenient means of acquiring the matte.

We show in this sketch that accurate environment mattes can be computed from natural images, without the need for specialized calibration of the acquisition. The goal is to take a set of example images, containing the optical element of interest (e.g. the lens in figure 1), and transfer the element’s environment matte to a new background image (example in figure 3).



Figure 1: **Input:** Three of a sequence of 42 images, static optical element (magnifying glass), moving background. The environment matte is computed using only the information in these images.

The technique is best understood by working backwards from the final composite of a novel background image  $N$  and the computed environment matte. Each pixel in the output collects light from a blend of pixels in  $N$ . Let us call the set of pixels which contribute to a given output pixel  $p$  the *footprint* of  $p$ , or  $p$ ’s *receptive field*. Previous researchers have defined the footprint using rectangular regions [Zongker *et al.* 1999] or mixtures of Gaussians [Chuang *et al.* 2000]. In this work, we must deal with complex multimodal distributions, so we use a discrete map of source pixels, where each source pixel has an associated weight. The value of the output pixel is then computed as a weighted sum over the pixels of  $N$ . Thus if we can compute the receptive field for each pixel, we can compute the composite.

In order to compute the receptive field of a given pixel  $p$ , we need at least two images: one containing the test object (e.g. the lens in figure 1), and one containing only the background (figure 2). We note that pixels in the background which have contributed to  $p$ ’s colour will have similar colour to  $p$ . In fact, for each background pixel, the similarity between its colour and the query colour is a function of the amount that background pixel contributes. Thus, we can obtain a bound on  $p$ ’s receptive field by computing the correlation between a small (e.g.  $3 \times 3$ ) window around  $p$  and each location in the background image. Such a bound is illustrated in figure 2c. Of course, for a single image, this bound is very weak—many pixels which accidentally share  $p$ ’s colour are included in the receptive field. However, with a sequence of images, as in figure 1, the receptive field is constant as the background moves, and with each new image, the footprint can be refined. Figure 2d shows the refined receptive field for the indicated foreground pixel after 8 views have

**Step 1:** Compute clean background for each image:



**Step 2:** Receptive field (RF) computed for one pixel. (a) Foreground. (b) Background. (c) RF for this pair. (d) RF from all pairs.



Figure 2: **Steps in the algorithm**



Figure 3: **Output:** Recovered environment matte over new image. Compare the environment matte (above red line) and transparency (below red line).

been integrated. Note how the single peak corresponds to the true source pixel, indicated in figure 2b.

Computing the background image may be achieved by mosaicing the moving-background sequence [Irani *et al.* 1994] or moving the camera. Figure 4 shows an example where the camera is moved to obtain a clean view of the background. In this example, there is just one reference view, so strong regularizing constraints were employed in order to permit a solution: the receptive fields were assumed small and close to their source pixels.



Figure 4: **Base image, single reference view of background, composite using computed environment matte.**

The examples show that, although its performance is scene-dependent, the technique can work well given sufficiently rich backgrounds, or sufficiently many images. They demonstrate that environment mattes can be captured under less stringent assumptions than have previously been described.

## References

- CHUANG, Y.-Y., ZONGKER, D. E., HINDORFF, J., CURLESS, B., SALESIN, D. H., AND SZELISKI, R. 2000. Environment matting extensions: Towards higher accuracy and real-time capture. In *Proceedings of ACM SIGGRAPH*, 12–130.
- IRANI, M., ROUSSO, B., AND PELEG, S. 1994. Computing occluding and transparent motions. *Intl. J. Computer Vision* 12, 1, 5–16.
- ZONGKER, D. E., WERNER, D. M., CURLESS, B., AND SALESIN, D. H. 1999. Environment matting and compositing. In *Proceedings of ACM SIGGRAPH*, 205–214.

\*e-mail: {wexler,awf,az}@robots.ox.ac.uk



# Image-Based Illumination for Electronic Display of Artistic Paintings

Da Young Ju<sup>\*</sup>, Jin-Ho Yoo<sup>\*</sup>, Kyoung Chin Seo<sup>\*</sup>, Gregory Sharp<sup>†</sup> and Sang Wook Lee<sup>\*</sup>

<sup>\*</sup>Department of Media Technology, Sogang University, Seoul, Korea

<sup>†</sup>Department of EECS, University of Michigan, Ann Arbor, Michigan, USA

## 1 Introduction

Visual impressions from two-dimensional artistic paintings greatly vary under different illumination conditions, but this effect has been largely overlooked in most poster productions and electronic display. The light-dependent impressions are more pronounced in oil paintings and they arise mainly from the non-diffuse specular reflectances. We present an efficient method of representing the variability of lighting conditions on artistic paintings utilizing both simple empirical reflectance models and an image-based lighting method. The Lambertian and Phong models account for a significant portion of image variations depending on illumination directions, and residual intensity and color variations that cannot be explained by the reflection models are processed in a manner that is similar to the image-based lighting methods. Our technique allows brush strokes and paint materials to be clearly visible with relatively low data dimensionality.

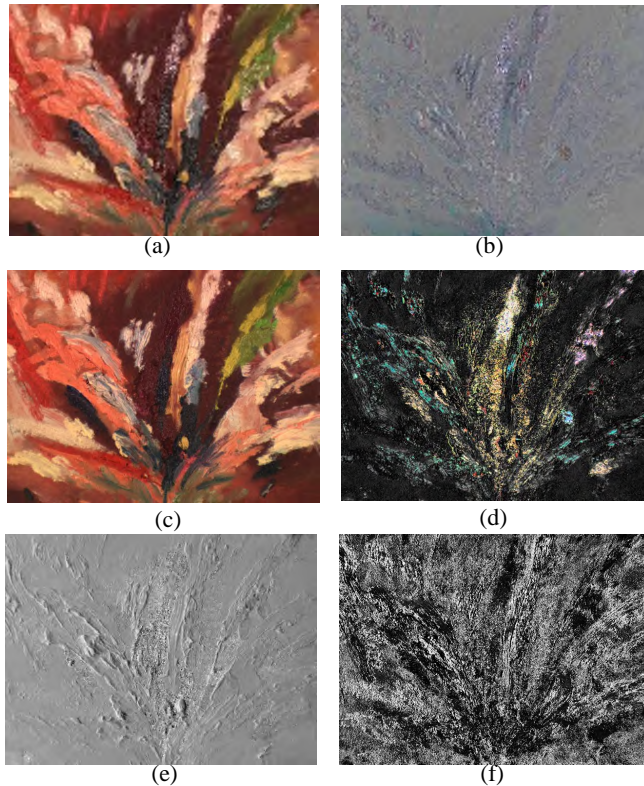


Figure 1. Oil painting: (a) Captured image  $I(x,y,\phi=90^\circ)$ , (b) Residual  $r(x,y,\phi=90^\circ)$ , (c) Diffuse  $C_d(x,y)$ , (d) Specular  $C_s(x,y)$ , (e) Surface structure  $\zeta(x,y)$ , (f) Surface shininess  $n(x,y)$ .

The visual perception of effects such as brush strokes and canvas texture to a large extent from the non-Lambertian specular reflections and the appearance of a painting (especially oil painting) varies greatly depending on the illumination conditions. A straightforward way to provide all the visual information would be to photograph a painting under all the possible illumination conditions and let a person change the conditions electronically in a computer. However, the data dimensionality would be prohibitively high, especially in a web environment.

## 2 Methods

We present a new strategy of efficiently displaying artistic paintings under a range of lighting conditions using a combination of input images and reflectance models. To explain this concept, we represent the image irradiance  $I$  as:

$$I(x,y,\phi) = I_d(x,y,\phi) + I_s(x,y,\phi) + r(x,y,\phi)$$

where  $I_d(x,y,\phi)$  and  $I_s(x,y,\phi)$  are the diffuse and specular reflections, respectively,  $\phi$  denotes the lighting angle, and the residual  $r(x,y,\phi)$  is the difference between the modeled image irradiance and measured image  $I(x,y,\phi)$ . The residual will be used to account for modeling error, self-shadowing, sub-surface scattering and interreflections. A Phong models is used to model the diffuse and specular components.

Nonlinear optimization is used to extract the most likely values for the diffuse and specular coefficients, from which the residuals are subsequently computed. The residuals are then compressed using principle components analysis. Since even the Phong model can account for a significant portion of lighting effects, the residual  $r(x,y,\phi)$  becomes very small, and amenable to high compression rates. Figure 1 shows an image of an oil painting captured with the illumination angle  $\phi=90^\circ$  in (a), its residual image in (b), and the relative values of the eight recovered diffuse and specular coefficients in (c)-(f). Figure 2 shows the compression rates achieved as a function of the number of residual image coefficients retained. An RMS error of less 4 grey levels is achieved using 13 coefficients.

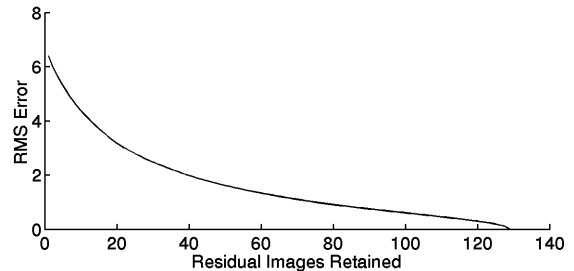


Figure 2. Reproduction accuracy as a function of the number of residual images retained.

See also: <http://www.sogang.ac.kr/~gsmc/svl.htm>

# Implementing the Continuous Staircase Illusion in OpenGL

Mark W. Scott  
mwscott@usit.net

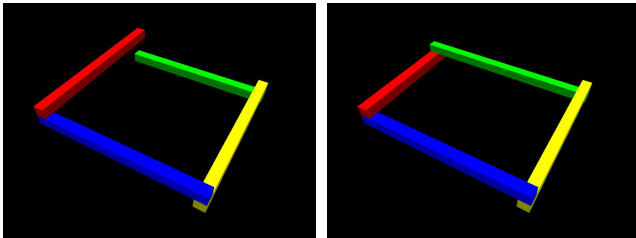
## Introduction

The inspiration for the project described in this sketch was certainly the M.C. Escher lithograph ASCENDING AND DESCENDING. An image of this drawing can be seen at <http://www.worldofescher.com/gallery>. It is based on the continuous staircase illusion of L. S. and Roger Penrose and depicts a three-dimensional scene that appears to have properties that contradict what is possible in an actual Cartesian representation of spatial objects. The goal of this project was to add animation and a navigable viewpoint to this illusion by using the three-dimensional environment of OpenGL.

## Attacking the Impossible

OpenGL is subject to the same Euclidean restraints that Escher's lithograph seems to violate, so it was clear that any degree of success would require deceptive modeling.

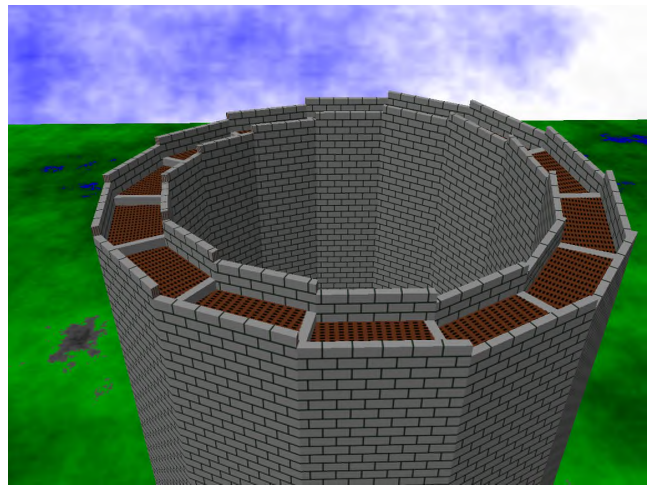
The first technique considered was to employ hidden gaps in the modeling. This trick is demonstrated in the next two screen shots. The first is non-deceptive and shows the vertical gap between the green and red bars. The second is derived from the first by extending the green bar and removing part of the red bar.



While this method produces a result similar to Escher's rendition, it presents many complications. It is not obvious how to position the gap in response to a changing viewpoint. Also, since it requires some of the sides' lengths to differ by a considerable amount, the problem of hiding model distortions as the point of observation is modified appears formidable.

Fortunately, an alternate technique was devised which is much simpler to implement in OpenGL. The inspiration for the new method is founded in the realization that the data reduction required to represent a three-dimensional model on a two-dimensional screen insures the existence of distinct polygons that share identical outlines to a viewer. In particular, it is possible to create sloped surfaces whose outlines will exactly match the outlines produced by horizontal surfaces. Using this tool, it is no longer necessary to have gaps in a continuous staircase model, since the elevation gains produced by climbing a step can be canceled by elevation loss in stair sections that only appear to be horizontal. OpenGL's ability to assign surface normals that are independent of a polygon's actual orientation even provides some extra support to the illusion that is not available in the real world. Several methods of adding slope to pseudo-horizontal surfaces were tested. The one chosen depends on computing the angle between the horizontal component of the viewing direction and the natural direction across the surface used when ascending the

stairs. The resulting slope will be directly proportional to the absolute value of this angle's cosine. The effectiveness of this method can be seen below in a screen shot from the current version of the OpenGL program.



## Animation

The animation included in the program consists of a ball that is allowed to bounce down the stairway. The physics simulator, used to control the ball motion, computes the ball's display position in two phases using two types of models. The first phase works with local models that have actual horizontal surfaces instead of pseudo-horizontal surfaces but are linked point by point to the model used for display. Realistic computations are made for the ball's movement in this local environment. The second phase finds the point in the local model directly under the computed position of the ball and its height above that point. The height is added to the linked point from the display model to get the display position for the ball.

Aids are included in the physics modeling to increase the probability that the ball can bounce repeatedly about the stairway. The walls are invisibly extended to keep the ball from escaping off the staircase and a slight nudge in the direction of the next step downward is included in each bounce.

## Possible Future Extensions

The mathematical modeling concepts exist for several potential enhancements that may be implemented when coding time can be allocated. These include the addition of shadows, alternate staircase designs that are less symmetric, some sort of ascending animation, and landscaping at the base of the tower that employs the same technique used in the staircase to create a continuous hill illusion.

As new features are added, the copy of the program available at [home.usit.net/~mwscott](http://home.usit.net/~mwscott) will be updated.



# Improving frameless rendering by focusing on change

Abhinav Dayal  
Dept. Computer Science  
Northwestern University  
abhinav@cs.nwu.edu

Benjamin Watson  
Dept. Computer Science  
Northwestern University  
watsonb@cs.nwu.edu

David Luebke  
Dept. Computer Science  
University of Virginia  
luebke@cs.virginia.edu

Realtime rendering requires accurate display of a dynamic scene with minimal delay. Frameless rendering [Bishop et al. 1994] offers unique flexibility in this regard: because it samples time per pixel, it can respond to change with very little delay, and at any location in the image. However, sampling is random, resulting in blurring in changing image regions. We present an approach for improving frameless rendering by making sampling sensitive to change in the image, as suggested in [Bishop et al. 1994]. By measuring this change in visual terms, we are able to direct sampling to those regions of change. The resulting algorithm produces sharper imagery, while introducing minimal overhead into the standard frameless algorithm.

We measure change by monitoring color differences in the image, using the summed squared difference between component colors at a pixel in the previous and current rendering. We use a probability distribution function (PDF) to choose the next pixel rendered so that changing image regions are sampled more frequently. The probability of every pixel is the weighted sum of its color difference and its age (time since it was last updated) both normalized over the entire image. The former biases rendering toward regions of change and the latter monitors for change in previously static image regions. This ensures that all pixels are sampled with a certain minimal frequency. The PDF is subsampled into rectangular tiles in image space. Besides bringing obvious improvements in speed, subsampling implements a spatially coherent response to change: if one pixel is changing, neighboring pixels are likely also changing. The probability that one of the pixels in a tile will be rendered is computed using the summed probabilities of the component pixels, normalized by the summed probabilities of all pixels in the image. In order to determine which pixel to render, we first select the tile according to the subsampled PDF, and then randomly select a pixel within that tile with bilinear interpolation of the surrounding tile probabilities just as in bilinear texture filtering.

Our renderer displays sharper imagery while using the same number of rays as a conventional frameless renderer. Figures 1 and 2 show corresponding frames of a video at a simulated rendering rate of 900,000 rays per second. In interactive use, our current renderer casts roughly two thirds as many rays per second as the standard renderer, but the resulting images are still sharper.

Our major goals in future research will be improving dynamic image quality, evaluating and tuning the performance vs. accuracy tradeoff, and comparing this approach to existing approaches. To improve image quality, we will investigate alternative image quality metrics, including sensitivity to spatial contrast and Gibsonian patterns of motion. We will also incorporate progressive rendering [Bergman et al. 1986] into the ray tracer, enabling control of the tradeoff between temporal and spatial accuracy – a line of research that we believe will be particularly fruitful. In evaluation and comparison, our standard will be a recorded animation that simulates the performance of a hypothetical instantaneous renderer, one that can produce frames in response to input without any delay. We can then compare the performance of our (or any) algorithm to the output of this hypothetically ideal renderer moment by moment. Ultimately, we

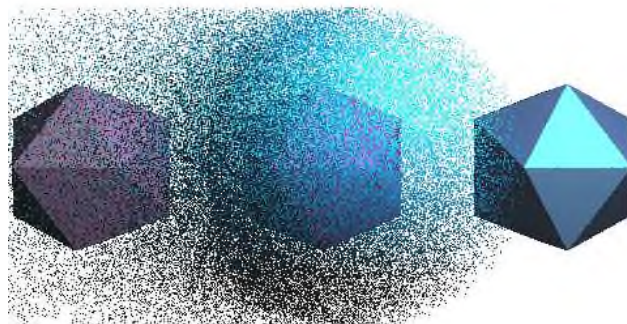
will examine the use of our change sensitive approach in the context of distributed rendering [Parker et al. 1999].

## Acknowledgements

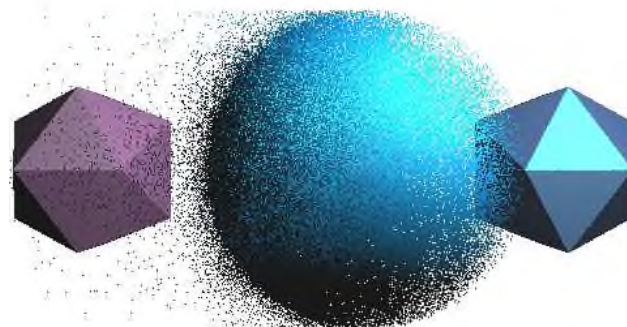
Our thanks to Conrad Albrecht for his late night coding efforts.

## References

- BISHOP, G., FUCHS, F., MCMILLAN, L. AND ZAGIER, E. 1994. Frameless Rendering: Double Buffering Considered Harmful. In *Proceedings of SIGGRAPH 1994*, ACM Press/ACM SIGGRAPH, New York. A. Glassner, Ed., Computer Graphics Proceedings, Annual Conference Series, ACM, 175-176.
- PARKER, S., MARTIN, W., SLOAN, P., SHIRLEY, P., SMITS, B. AND HANSEN, C. 1999. Interactive Ray Tracing. In *Proceedings of ACM Symposium on Interactive 3D Computer Graphics*, 119-126.
- BERGMAN, L., FUCHS, H., GRANT, E. AND SPACH, S. 1986. Image Rendering by Adaptive Refinement. In *Proceedings of SIGGRAPH 1986*, *Computer Graphics* 20, 4, 29-37.



**Figure 1:** Traditional fully random frameless rendering. A sphere moves across quite visible background objects.



**Figure 2:** A frameless renderer that responds to change. The same sphere now occludes the background with fewer rays.

# Integrating Lenticular into Digital Printmaking

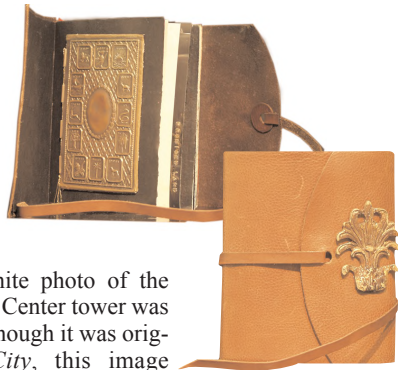
Dorothy Simpson Krause\*  
 Professor Emeritus, Massachusetts College of Art



*Promised Land* 16x24 inches (40 x 60 cm) ©Krause 2002

On September 11 I was in the Middle East on my continuing quest to understand why, in the name of religion, one would kill another who did not share their beliefs. The evening before I had crossed the border to Israel from Jordan, where I had photographed the crumbling sandstone ruins of Petra.

The collaged journal pages silhouetting the gilded city of Jerusalem against small pieces of gold and silver prayer paper were completed September 11, 2001. The following day, the black and white photo of the imploding World Trade Center tower was added to the image. Although it was originally called *Holy City*, this image became the symbol of the journal and the series. I changed its name to *Promised Land*.

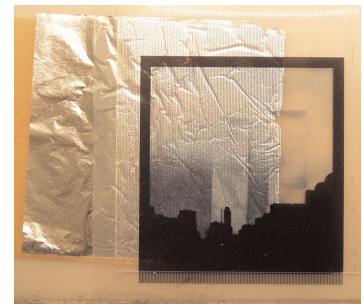
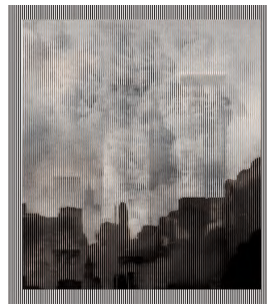
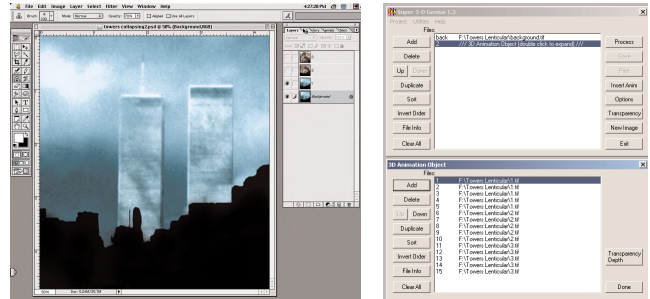


The two-page spread was printed using the Wasatch RIP, on the Epson 9500 on Hahnemuhle Torchon fine art paper.

Using a photograph taken by Swiss photographer, Viola Kaumlen, I made a small lenticular image of the tower imploding. The first image shows the towers prior to the impact; the second shows them collapsing and the third shows the clouds of ash.



In Adobe PhotoShop the three images were aligned and each layer saved as a separate tiff file. Using the lenticular software, 3D Genius, the tiff files were put into an animation, cut and reassembled in vertical strips. This interlaced file was saved as a tiff file and printed with the Epson 9500 on clear film.



It was carefully aligned behind a 30-line lens, a sheet of plastic with a series of parallel lens or lenticules embossed into the surface, so that only one frame is seen at a time. Silver leaf was put behind the lenticular.

A rectangle the size and position of the lenticular was cut out of a piece of 16" x 24" matte board. The print was glued to the matte board and the printed tower image cut away leaving a hole. The small lenticular was fitted into the space, level with the surface. For added reflectivity, silver leaf was added around the lenticular.



As you move past the image, the collapse is continuously "replayed" until you stop, be still and focus on the part of the image you choose to see. It is that ability to view selectively that enables to go forward in our lives and in our art.

\*dotkrause@dotkrause.com



# Interactive Level-Set Tools for Photo Editing

Thomas Howard and Bryan Morse\*  
Department of Computer Science, Brigham Young University

## 1 Introduction

This sketch presents a suite of interactive image-editing tools based on properties of and manipulation of image level sets. This suite includes level-set smoothing, level-set constrained sharpening, and level-set “nudging” (image distortion).

The smoothing tool preserves edge sharpness while smoothing noise and jagged contours. The sharpening tool, constrained by local level-curve properties, is less sensitive to noise than standard techniques. The nudging tool allows incremental movement of image content without the artifacts and overhead of global mesh-based distortion techniques.

All of these tools use partial differential equations to change the image content, but unlike existing approaches using PDEs, the location and duration of application are controlled interactively with immediate feedback to the user.

## 2 Tools

A *level set* is the set of all points in an image sharing a specific value. Although discrete pixels may not share values locally, we may assume continuity of the image intensity function and infer the existence of intermediate values. Thus, these level sets form closed curves in the image plane.

In this way, one can think of the image not as an array of pixels but as a *topographic map*. Properties of the local contour can be used to drive or control other operations. Moreover, we can directly manipulate the contours themselves—smooth them, move them as desired, etc.

The interactive level-set smoothing tool is based on established techniques for level-set smoothing [Sethian 1996; Alvarez et al. 92; Morse and Schwartzald 2001], with the following contributions:

- Interactive user control of the target area and the amount of smoothing performed,
- Automatic parameter estimation based on local image characteristics (no trial-and-error tweaking of the smoothing parameters), and
- Time-varying adjustment of smoothing parameters based on the duration of the user’s application of the tool.

The interactive level-set nudging tool is also based on established techniques for manipulating level sets [Sethian 1996], but unlike prior applications, specification of the local motion vector field is controlled by the user. Unlike mesh-based deformation approaches, the effect on the image is *local* and *direct*. Thus, subtle movement of image content can be interleaved with smoothing, sharpening, or any other desired interactive operations. This sets it apart from the range of existing mesh-based deformation tools currently available.

Unlike the smoothing and nudging tools, the level-set constrained sharpening tool does not directly manipulate level-set contours. Rather, it uses the local level-set curvature as a constraint to prevent oversharpening (accentuation of noise and corners). In this way, it responds better in the presence of noise than current interactive sharpening techniques.



Figure 1: Nudging tool used to produce a distorted-mirror effect



Figure 2: Simulated weight loss using nudging and smoothing

## 3 Conclusion

These tools demonstrate the power of locally-applied PDEs under user control. In addition to the operations incorporated into this suite of tools, there are many other PDE-based algorithms for image enhancement [Sapiro 2001, for an excellent survey], which we believe might be successfully incorporated into locally-applied interactive tools.

## References

- ALVAREZ, L., GUICHARD, F., LIONS, P.-L., AND MOREL, J.-M. 92. Axioms and fundamental equations of image processing. Tech. Rep. 9231, Ceremade, Université Paris-Dauphine, Place de Lattre de Tassigny, 75775 Paris Cedex 16 France.
- MORSE, B. S., AND SCHWARTZALD, D. 2001. Level-set image reconstruction. In *Proceedings Computer Vision and Pattern Recognition (CVPR)*, IEEE Computer Society Press.
- SAPIRO, G. 2001. *Geometric Partial Differential Equations and Image Analysis*. Cambridge University Press.
- SETHIAN, J. A. 1996. *Level Set Methods*. Cambridge University Press.

\*e-mail: morse@cs.byu.edu

# Interactive Visualization of Large Scale Time Varying Data Sets

Patric Ljung, Mark Dieckmann, Anders Ynnerman  
Scientific Visualization Group  
University of Linköping  
Campus Norrköping, Sweden  
{plg,mardi,andy}@itn.liu.se

Visual analysis of time varying scientific data can be divided into four different categories with an increasing degree of user interaction.

- 1) Production of static images representing scientific data at selected times.
- 2) Production of video sequences in which graphical representation, time line and viewpoints are predefined.
- 3) Interactive streaming of logged data sets, allowing the user to alter graphical representation, filtering, time lines and viewpoints.
- 4) Real time interaction with the simulation or experiment that produces the data, allowing the user to alter parameters, graphical representation, filtering, time lines and viewpoints.

Research on tools for interactive streaming of large data sets (category 3) have a potentially very large impact on both the end usage and visualization research as such. As data sizes grow visualization of this kind will become extremely valuable and in some cases essential for data analysis. Research efforts in the past have, however, to a large extent focused on categories 1, 2 and 4. Category 3 visualization has thus not been developed to the point where it can meet the increasing demand from application areas. It thus constitutes a field with a number of challenging research issues that need to be addressed in order to develop the needed tools. These tools constitute a novel approach to the visual analysis of data sets that will enable new approaches to efficient data post-processing and analysis. The research focus in this project lies on the issues encountered when developing analysis tools within the framework of category 3 visualization. The goal of the project is to develop underlying methods, which enable the implementation of a visualization system for large-scale time-varying data. The key demands put on the user level features are:

- Arbitrary play back speed and direction.
- Sustained frame rate and/or rendering quality.
- Full 3D navigation and manipulation in real-time.
- Flexibility with respect to the type of visualized data.
- Implementations on low end desktops as well as high end immersive VR environments.

In a prototype software implementation we have used streaming of a time varying data set from a kinetic plasma simulation. The implementation facilitated the streaming of particle data and rendering of point primitives using a brute force approach. We have also used a pre-binned and size limited data set to experiment with real-time volume rendering based on 3D textures. The prototype software was applied to a case study in which a study of the time evolution of a simulation of an ion beam instability as found in supernova remnant shocks was used McClements et al. [2001]. In figures 1 and 2, we show two snap shots with different views, at the same time step, of the evolved plasma instability. The results are very promising and early results were reported on in Ljung et al. [2000] and Dieckmann et al. [2002].

In our current work we are extending the project to also treat time varying datasets with attribute data of higher dimensionality such as vector and tensor fields. Hardware constraints have so far forced

us to consider only a subset of the simulation particles for the particle visualization or a limited resolution of phase space for volume rendering.

We are, however, now aiming at the combination of both methods using hybrid techniques to provide information about individual particle trajectories and at the same time visualize particle densities with the full accuracy of the simulation. We are also addressing the issue of adaptive (multiresolution) visualization. The current work is based on our previously developed data management system.

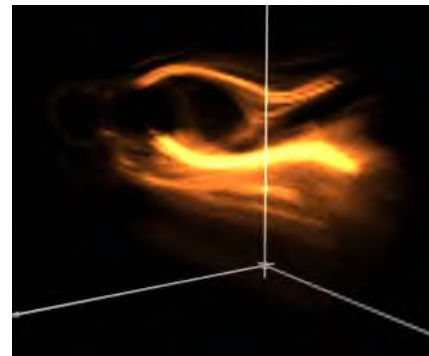


Figure 1: Electron phase-space distribution at viewpoint 1.

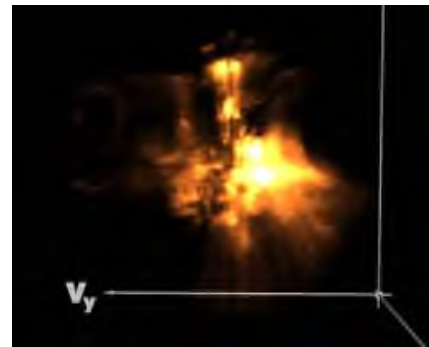


Figure 2: Electron phase-space distribution at viewpoint 2.

## References

- MCCLEMENTS, K.G., DIECKMANN, M.E., YNNERMAN, A., CHAPMAN, S.C. & DENDY, R.O. 2001. Surfatron and stochastic acceleration of electrons at supernova remnant shocks, In *Phys. Rev. Lett.*, 25, No. 255002.
- LJUNG, P., DIECKMANN, M., ANDERSSON, N. & YNNERMAN, A. 2000. Interactive Visualization of Particle In Cell Simulations. In *Proceedings of IEEE Visualization 2000*, pages 469-472.
- DIECKMANN, M. E., LJUNG, P., YNNERMAN, A., MCCLEMENTS, K.G. 2002. Three-dimensional visualization of electron acceleration in a magnetized plasma, In *IEEE Transactions on Plasma Science*, accepted for publication.



# Interruptible Rendering

J. Cliff Woolley  
University of Virginia  
jwoolley@virginia.edu

David Luebke  
University of Virginia  
luebke@virginia.edu

Ben Watson  
Northwestern University  
watsonb@cs.nwu.edu

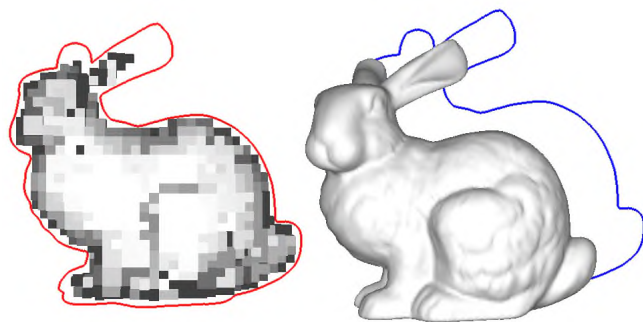
## Abstract

Interruptible rendering is a novel approach to the fidelity-versus-performance tradeoff ubiquitous in real-time rendering. Interruptible rendering unifies *spatial error*, caused by rendering coarse approximations for speed, and *temporal error*, caused by the delay imposed by rendering, into a single image-space error metric. The heart of this approach is a progressive rendering framework that renders a coarse image into the back buffer and continuously refines it, while tracking the temporal error. When the temporal error exceeds the spatial error caused by coarse rendering, further refinement is pointless and the image is displayed. We discuss the requirements for a rendering algorithm to be suitable for interruptible use, and describe one such algorithm based on hierarchical splatting. Interruptible rendering provides a low-latency, self-tuning approach to interactive rendering. Interestingly, it also leads to a “one-and-a-half buffered” approach that renders sometimes to the back buffer and sometimes to the front buffer.

## 1 Overview

Computer graphics practitioners have long recognized the tradeoff between spatial detail and frame rates. Given additional resources, which is more important: more detailed models or higher frame rates? Answers to date have been ad hoc. Conventional wisdom simply dictates that high, constant frame rates are best; applications tend to target 30-60 Hz frame rates.

*Interruptible rendering* provides a principled approach to this tradeoff by unifying spatial and temporal error in a single metric: screenspace distance (Figure 1). A coarse image is generated and continuously refined, decreasing spatial error while temporal error grows with the passage of time. When temporal error exceeds spatial error, there is no longer any reason to refine further: any improvement to the appearance of objects in the image will be overwhelmed by their wrong position and/or size. In other words, when the error due to the image being *late* is greater than the error due to the image being *coarse*, taking time for further refinement is pointless. The front and back buffers are then swapped and rendering begins again into the back buffer for the most recent



**Figure 1: Spatial and temporal error.** The ideal instantaneous image that reflects the most up-to-date input is shown in silhouette (colored outlines). The left image is coarsely sampled, representing some spatial error. The right image is finely sampled, but as a result is quite late. The resulting temporal error, manifested as image-space distance from the ideal instantaneous image, is larger than the spatial error of the coarse approximation. In this sense the coarsely sampled bunny actually represents lower visual error.



**Figure 2:** A sequence of progressively refined, splatted volumetric models. Each finer model contains and covers all coarser models.

viewpoint. A system which minimizes combined spatial-temporal error quite intuitively results in coarse, high frame rate display when input is changing rapidly, and finely detailed, low frame rate display when input is static.

**Interruptible Image Generation:** As the name implies, interruptible rendering requires an interruptible image generation process – interruptible because a sudden motion by the user can drive up temporal error at any time. The back buffer should always contain a complete image ready to be displayed, with the rendering process incrementally refining the image until further refinement is pointless. This requires an interruptible, progressive image generation process. Depth buffered rendering requires a continuous LOD algorithm that preserves *containment* – refined versions of an object completely contain more simplified versions already rendered. We use a hierarchical splatting algorithm similar to QSplat [Rusinkiewicz 2000], but using *bounded* rather than *unbounded* volumes to ensure containment (Figure 2). Other possibilities include Sander’s *progressive hull* structure [Sander 2000] or adaptive ray tracing.

**Measuring Temporal Error:** We currently measure temporal error by tracking the corners of an object’s bounding box to quickly estimate how far it may have moved since rendering began on the current image. For a more accurate, scalable approach we plan to investigate randomized algorithms that avoid bias by continuously picking a random set of vertices to track from the model.

## 2 Benefits and Implications

Interruptible rendering is by nature a low-latency algorithm: temporal error is continuously monitored (we check once per millisecond in our current implementation), so the system can respond almost instantly if the user makes a sudden move that invalidates the current image. By addressing temporal error, interruptible rendering also intrinsically tunes spatial detail to the rendering platform performance. On a low-end platform, less detail can be drawn before temporal error dominates spatial error and a new image is started. Thus when input is changing rapidly, the system will render less detail than a high-end platform but be just as responsive. Finally, an interesting corollary of unifying spatial and temporal error is that sometimes the image being refined in the back buffer may achieve lower combined error than the image in the front buffer. This implies that we should not continue to display the front image, but instead should swap buffers and continue refining the image *in the front buffer*.

## References

- RUSINKIEWICZ, S., AND LEVOY, M. 2000. QSplat: A Multiresolution Point Rendering System for Large Meshes. In *Proceedings of ACM SIGGRAPH 2000*.
- SANDER, P., GU, X., GORTLER, S., HOPPE, H., SNYDER, J. 2000. Silhouette Clipping. In *Proceedings of ACM SIGGRAPH 2000*.

# Inverse Texture Warping

Robert Falco  
The Secret Lab  
robert.falco@disney.com

Hank Driskill  
The Secret Lab  
hank.driskill@disney.com

## 1 Introduction

Standard texturing of a geometric surface works by mapping a 2D image into the parametric space of the surface. This technique works great whenever the texture image is created (painted, etc) for a specific surface with predetermined contours. Serious problems arise when the same texture is applied to the same surface but after some deformation has been applied. The result is the all too familiar rubbery texture look. This stretchy appearance can be particularly noticeable on surfaces that have identifiable traits such as scars, holes, text, repeating patterns, or in our case reptilian scales.

This sketch describes a technique used on “Reign of Fire” to minimize and isolate the ugly texture stretching that occurs when a deforming surface gets the same texture image applied over successive frames of deformation. The technique was specifically developed to allow the alligator-like skin features on the faces of our dragons to “stick” on the skin surface while the areas between the features were allowed to stretch.

Production constraints helped to narrow the list of possible technical solutions and guided the final workflow. For example, since production had already begun, we were unable to alter existing models or animation. Impacts on artist, disk space, and render resources needed to be kept to a minimum. The additional overhead meant limiting use to a select few close-up shots where the added level of quality would be visible.

## 2 Concept

Given a reference surface (RS), a reference texture image (RT), and a deformed version of the reference surface (RS'), generate a new texture image (RT') such that when RT' is mapped onto RS' specific areas of the reference texture image do not appear to undergo stretching. In other words, make it so important areas of a surface (a hole, a scar, etc) don't appear to stretch while other portions of the surface take up the slack. Determine the deformation that is occurring on the surface and apply the exact opposite to the texture image in order to “undo” any warping.

## 3 Process

The technique is fairly easy to implement and fits conveniently into most production pipelines. The process is as follows:

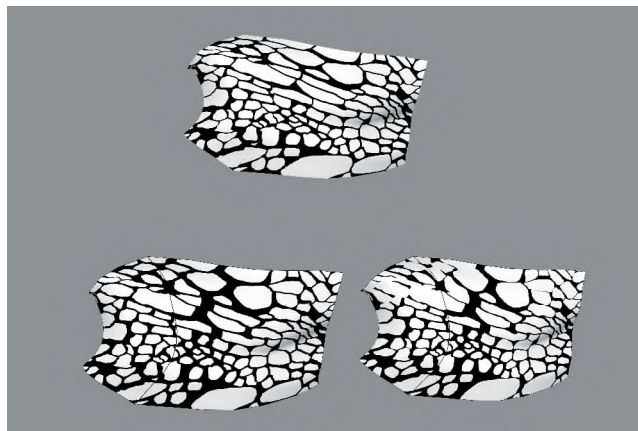
- Identify areas that are allowed to stretch and areas to remain rigid. Luckily, the dragons in “Reign of Fire” already had high contrast displacement maps used to indicate reptilian surface features. Apply object detection software on the images and store all information on feature centers and vectorized contours. These features will define the areas to minimize texture warping. This step is independent of the actual rendering process and only needs to be done once.

- At render time, sample RS and RS' at regular intervals in parameter space. For each sample calculate the magnitude of  $\Delta\partial P/\partial u$  and  $\Delta\partial P/\partial v$  between the two surfaces. This can be thought of as how much RS' has expanded or contracted within the sample area relative to the reference surface. This step only needs to be done once per surface per frame.

- All feature contour points need to be moved relative to the amount of deformation occurring on the surface, but in a direction opposite of the deformation. This is done by summing all the deformation deltas encountered when going from the contour point to the feature center. This summation will indicate either an expansion or a contraction from the original distance. Move feature points the summed amount but in the opposite direction. For simplicity we only applied deltas which indicated an occurrence of expansion. Handling compression would mean dealing with the implications of colliding features. It was also necessary to constrain feature points that land on image borders in order to maintain coherency across patch boundaries.

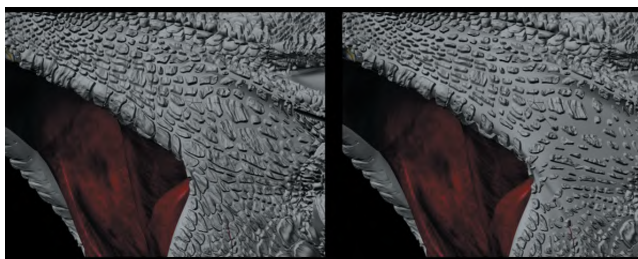
- Create a single 2D constrained Delaunay polygon mesh from all the new feature contour points. Each new contour point should carry along the original uv coordinate of its location within RT. Texture map the polygonal mesh onto a flat plane. The resulting image is your new texture image and should be used during final rendering in place of the original reference texture.

Figure A



Top: standard texturing  
Left: adjusted for expansion only  
Right: adjusted for expansion & compression

Figure B



Normal texturing:  
stretching occurs  
Adjusted texturing:  
scales maintain shape

# Investigating Face Space

Steve DiPaola  
Simon Fraser University  
steve@dipaola.org

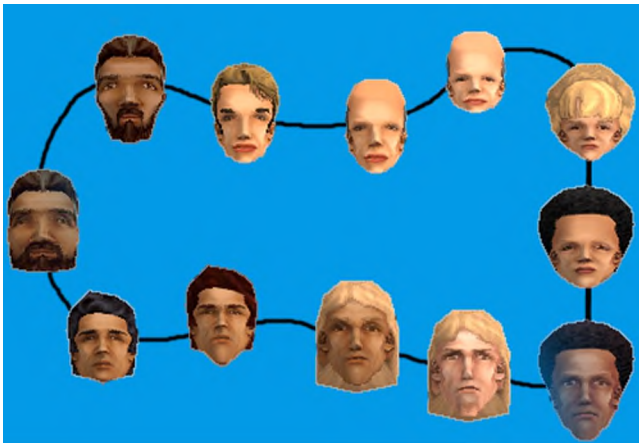
## 1 Introduction

Imagine an n-dimensional space describing every conceivable humanoid face, where each dimension represents a different facial characteristic. Within this continuous space, it would be possible to traverse a path from any face to any other face, morphing through locally similar faces along that path.

We will describe and demonstrate a development system we have created to explore what it means to “surf” through face space. We will present our investigation of the relationships between facial types and how this understanding can be used to create new communication and expression systems.

## 2 The System

While a multi-dimensional space is a convenient way to generate a universe of faces, it has a slippery relationship to the ways we perceive facial similarity and sort faces into different categories. We currently measure facial similarity by values of low level facial parameters, such as chin width, eye radius, etc., based on our knowledge of creating facial animation systems for the graphics and avatar industries [DiPaola 1991; Parke et al. 1996], however perceived similarity operates on a different metric. The correlation between facial parameter values and perceived race, age, gender, or cultural stereotype is even less straightforward. It is this experimentation with classifications and their facial relationships, as well as how a viewer reacts to them, which have provided the motivation to create this development environment.



The system uses a genetic algorithm technique known as aesthetic selection [Sims 1991] to search through face space, using approximately 25 parameters to specify the genotype. An earlier version of this system was developed as a freely downloadable web tool, which allows users to create families of 3d heads for use in the popular simulation game “The Sims” by Maxis. In this version a menu of random faces is generated from random points in face space. By selecting one of the faces, the user is able to surf through the local region centered around the selected face and

can increase or decrease the local region of similarly looking faces by the use of a mutation rate slider and other controls. In this way, novice users can “find” or breed a face or family of related faces that appeal to them for use in the game.

A large web based community is currently using this system to make thousands of faces. In addition, these users have the ability to publish and trade faces on the web and even breed their creations with those of other users. Every head or family of heads becomes a local starting point or region in face space. What we have found is that a user will have a strong emotional response to certain faces they encounter (which appears to have cultural underpinnings) and that this emotional attachment continues as they surf that locally related face space. By expanding this system into a full development tool, we hope to better understand these cultural relationships for both specific faces as well as their relationship to other faces in face space. Specifically, we are interested in the following areas:

- The system allows us to traverse morphing face [Blanz et al. 1999] paths through face space, with the ability to affect the trajectory of path (straight, curved) and the key faces to target along the way. Music can be described as paths through tonal space with significant cultural biases described by certain paths or key points (i.e. notes, intervals, scales) determine whether certain motifs sound appealing. Can we similarly write a composition in face space? Imagine a representation of two intertwining paths through face space that use their relationships, both harmonic and discordant, to create a composition.
- We have so far concentrated on facial types, but we can expand the current development system to work with facial emotions as well. Can we create a syntax or language using face symbology that can be recognized rapidly, and span national and cultural boundaries?

It is well known that we humans have specific neuro-programming for recognizing and interpreting faces. It is hoped that we can use the face space development system to better understand the conscious and intuitive meaning of faces and the universal language they appear to represent (facial meaning and inter-relationships) for use in the arts, cultural theory and communication.

Acknowledgements: Roger Critchlow and David Collins for co-authoring the system; John Wentworth, Ai Ebtekar and Will Wright for their ideas and support. See: <http://www.dipaola.org/facespace>

## References

- DiPaola, 1991. Extending the Range of Facial Types, *Journal of Visualization and Computer Animation*.
- Parke, Waters, 1996. *Computer Facial Animation*. 101-4, 214-19.
- Sims, 1992. Artificial Evolution for Computer Graphics, *ACM Siggraph*.
- Blanz, Vetter, 1999. A Morphable Model For The Synthesis of 3D Faces, *ACM Siggraph*.



# Lewis the Robotic Photographer

Cindy M. Grimm\*      William D. Smart\*  
Department of Computer Science  
Washington University in St. Louis  
St. Louis, MO 63105  
United States

## 1 Introduction

Lewis is a (short) human-sized mobile robot that wanders through crowded rooms taking pictures of people, much like a photographer at a wedding reception does. The goal is to take high-quality, well-composed photographs of people non-intrusively, and to offer these pictures as keep-sakes of the conference.

## 2 Technical Challenges and Solutions

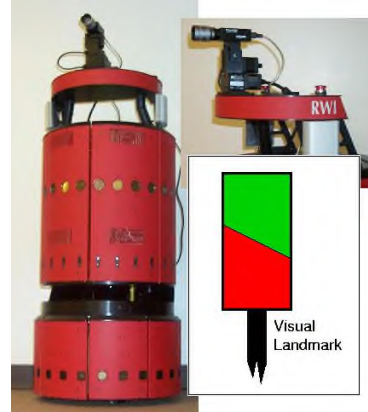
The main technical challenges for this installation arise from the sometimes conflicting goals of finding and framing good camera shots, and interacting safely with large crowds of people. To find good potential shots, we first attempt to find candidate faces in an image. These faces are then correlated with readings from the on-board laser range-finder to calculate the 2d floor position of each person. These positions are then used to calculate good positions from which to take a picture. Once a good position is found, the robot must plan a path to that point, possibly altering it on the fly to avoid people who might get in its way. Once the picture is taken, it is sent to an off-board workstation for further processing, selection and printing. During this process, the robot must constantly be aware of its position relative to the bounds of the installation, and interact safely with the conference attendees that it encounters. We briefly discuss our solutions to each of these key challenges below.

### 2.1 Finding Faces

We find candidate faces in a camera image by identifying contiguous regions of skin-colored pixels. All human skin falls into a (roughly) elliptical area in the HSV color space. By looking for contiguous areas of pixels with colors in this region, we can find skin-colored blobs in the image. We remove candidate regions that are the wrong shape (faces are roughly elliptical), or the wrong size (skin-colored walls, for example, are too large). We can also make simplifying assumptions, such as only looking for faces above a certain height, that reduce the amount of computation that we have to do.

### 2.2 Finding People and Calculating Positions

Once we have identified candidate faces in an image, we try to relate them to information from the on-board laser range-finder. This sensor returns 180 readings of radial distance from the robot, in one degree intervals, 90 degrees to each side of the front of the robot. We assume that each face has a corresponding pair of legs, and that these legs are more-or-less directly beneath the face. Since we know the correspondence between the data from the laser range-finder, we can associate the data from the laser with the candidate faces to determine the likely  $xy$  position of the owners of the faces. From these positions, we infer the direction in which they are looking (two people talking tend to look at each other), and calculate



a likely spot from which to take a good picture. Note that several such positions can be calculated from a single image, since there may be many potential shots. These positions are then passed to the path planner.

### 2.3 Path Planning and People Avoidance

Lewis uses standard mobile robot path-planning algorithms to plan trajectories to all of the candidate picture-taking positions. These trajectories are then ranked according to distance and difficulty, and the best one is chosen and executed. This process runs continually, constantly adding new goal points, and re-evaluating the paths to them. This allows Lewis to opportunistically change its mind about which picture to pursue. People are not static and tend to move about. This means that shots identified in the past will not be valid forever. A time decay mechanism is implemented that favors more recently added potential shots, culling out those that probably do not exist any more.

An important component of path-planning is staying within the bounds of the installation. Lewis uses internal odometry and a visual landmark to determine its exact location, which is then compared to a supplied map. The landmark is a cylinder with a diagonal stripe running around it, mounted on top of a post (see figure). By using the size, angular height and apparent proportions of the two colors, Lewis can reliably determine where in the installation it is.

## 3 Summary

This installation is essentially an integration project. Many of the individual components can be made to work in isolation. However, trying to move to a given position to take a picture, while trying not to stray too far from the installation, and making sure not to run over anyone's toes turns out to be surprisingly difficult. In addition to amusing SIGGRAPH attendees, our intention is to use this installation to test different interaction and localization strategies in a real-world environment, and use this information to build the next version of the system.

\*Email: {lewis,cmg,wds}@cs.wustl.edu



# “Loops” — a digital portrait

Marc Downie

The Synthetic Characters Group, MIT Media Lab, marcd@media.mit.edu

## Introduction

This sketch outlines some of the background to a collaborative artwork created in September 2001 by Marc Downie, Shelley Eshkar and Paul Kaiser. This digital portrait of dance legend Merce Cunningham uses as a point of departure a motion-captured recording of ‘Loops’, his solo dance for hands and fingers; it uses new real-time non-photorealistic rendering techniques; and it exploits an advanced behavior architecture to structure the performance of the piece. In the resulting animation motion-captured joints become nodes in a network that sets them into fluctuating relationships with one another, at times suggesting the hands underlying them, but more often depicting complex cat’s-cradle variations.

## Research Context

This work was based in part on research conducted by the MIT Media Lab’s Synthetic Characters Group — a group focused on the creation of autonomous synthetic creatures. These artificial intelligences inspired by biological systems typically ‘live’ in real-time graphical worlds and are capable of taking decisions based on their synthetic perceptions, their motor control, and their previous experiences [Blumberg et al. 2002]. These creatures are capable of a variety of adaptations to their virtual environment, possess non-representational bodies and are often rendered using a variety of non-photorealistic rendering techniques. Such creatures have in the past sought to unite research into synthetic perception, computer learning and dynamic visual form and have appeared in a number of SIGGRAPH installations (for example, [Downie 2001]). The Loops piece’s technology is based on this core conceptual engineering.

## Behavior

In the system that drives this piece, the motion-captured animation material drives the movement of 42 autonomous creatures. These creatures probabilistically take decisions concerning their appearance, the quality of their movement, and their structural connections to other points in the hands; they perceive other creatures in their environment and choose to send signals to them; they autonomously build expectations about these signals and react when these expectations are violated.

Using this system, the authors were able to assemble the piece rapidly and above all interactively. Since the focus of the technology used here has been to create interactive creatures, the collaboration could take place around a running system which could be manipulated and reconfigured in real-time. By using a simple network protocol the artists could move from perturbing a creature’s motivational variables, to manipulating its perception system, or adding new behaviors to its repertoire. Within this framework the collaboration began around the creation of a basic vocabulary of forms, movements, topologies and interactions for the creatures, and proceeded towards the creation of a ‘script’ that shapes the probabilistic action selections that the simple creatures make.

## Rendering

The distinctive graphical styles of these creatures’ bodies are created by two techniques: a computationally inexpensive approximation of ‘motion blur’ and the repeated overlapping of translucent, randomly perturbed geometry. Meshes used for the body of a creature are distorted to follow splines describing the connections that the creature has decided to create. Random perturbation around these curves is controlled by an interpolated set of example meshes created by hand. Perturbed in this fashion



and fused by the overlapping images from the motion blur, the bodies of the creatures can assume a wide range of forms from hand-drawn filaments, chalk-like marks, or amoeba-like jellies to solid planes in space. More images from this piece are available online [Kaiser 2001].

## Current Work

The piece premiered with some 40 basis behaviors in each creature and 11 minutes of script. The piece incorporates 12 minutes of motion capture material, 15 minutes of recorded narration by Cunningham, and 10 minutes of music by Takehisa Kosugi, and all these elements are looped indefinitely. Finally, the motion of the creatures is perturbed by the motion of people in the gallery space. The piece therefore unfolds across a number of time scales, with rendering parameters that change with every rendered frame, behaviors that develop over minutes, adaptations that occur over the course of an hour, and moments that occur perhaps only once a day.

In a wider context this work seeks to solve a now classic digital art problem — how to create a workplace within an exploration vs. exploitation trade-off. Can we create a system and an artistic process that is complex enough to contain surprises, but controllable enough to let artists take advantage of them when they occur? We feel that the use of the synthetic creature as metaphor, as research and as architecture is central to our success here in this piece.

The process described in this sketch is being extended to create projections for a new Cunningham piece for theater premiering in September 2002. To this end the behavior and learning capabilities of the creatures will be extended, the class of data that the creatures can respond to will expand (to include images and animation) and several new ways of interacting with the authorship environment (including a new behavior specification language) will be incorporated into the creatures.

## References

BLUMBERG, B., DOWNIE, M., IVANOV, Y., BERLIN, M., JOHNSON M. P., AND TOMLINSON, B. 2002. Integrated Learning for Interactive Synthetic Characters. In: *Proceedings of ACM SIGGRAPH 2002*, ACM Press / ACM SIGGRAPH, New York. Computer Graphics Proceedings, Annual Conference Series, ACM.

DOWNIE, M. 2001. music creatures. In: *Proceedings of SIGGRAPH 2001: Conference Abstracts and Applications* (Art Gallery).

KAISER, P. online at: <http://www.riverbed.com>

## Reality

The need to recreate reality is a driving force behind art. Mimesis – imitation of the real world as a main function of art was observed and analyzed as early as the fourth century B.C. by Socrates, Plato, and Aristotle. Although artists searched for inner inspiration, the psychological need to record reality was always present, appearing during the renaissance, in baroque painting, the art of film, and probably reaching it's most sophisticated form in computer graphics and animation. Seventy years ago, Polish critic Karol Irzykowski, in his book “The 10<sup>th</sup> Muse”, predicted that one day animated film would evolve into the most important film genre of the future. He also contemplated the possibility of animation reaching a state of realism in which it could show “ordinary things and people”. He would be surprised to learn how true his words were to become. Digital media has given us an opportunity to go even further towards that goal of “registering reality”. In this sketch, I present these concepts in relation to Digital Domain’s work on New Line Cinema’s release of J.R.R. Tolkien’s classic tale “Lord of the Rings: Fellowship of the Ring”.

## To kill a wraith

We were tasked with the animation of the imposing and ultimately drowned Ringwraiths in the magical flood sequence entitled “Ford of Bruinen”. Since, as animators, we aim to create the “Illusion of Reality”, our main objective was to create a realistic vision of the gruesome death of these characters.

Although from the very beginning we knew that our characters would play a secondary role to the digital water simulation in the sequence, we still decided to put as much detail as possible into them. We hope that the audience sensed the difference. The process started with character setup, designed to operate within Digital Domain’s proprietary character pipeline. Both the wraith and his horse were built with animatable proxy and high-resolution weighted models in addition to layers of cloth-simulated elements. Weighting of the high-resolution horses was done with extensive use of influence objects that represented major group muscles. Strategically placed, simple NURBS surfaces gave us satisfying default weighting results that sped up the exhausting process of binding the model to the skeleton. Since after each pass of animation we had to simulate cloth motion, it was really important to speed up that simulation process. Layers with cloth simulated elements contained simplified versions of character’s geometry together with collision objects representing the volume of water.

Computer animation has the same principles as any other animation technique. One could easily distinguish keyframes and breakdowns. What makes a difference is the use of computer-generated interpolation. While generated interpolation is a great way to visualize motion in it's simplest form, what the computer is not able to emulate is the “creative imperfection” of hand drawn or hand posed in-between frames. Realistic motion is not perfect; it’s full of inconsequence and expression. That is why extended use computer-generated interpolation often detracts from the movement of life. The fast motion that we dealt with in the flood sequence was especially demanding. Just as a traditional cell animator would hesitate to give their in-betweeners just a couple of keyframes for a sequence containing high-speed and complex motion, we hesitated with the computer. Thus, our animation of the characters in this sequence entailed many more detailed

keyframes then is typically warranted in computer animation. In this regard, Alias Wavefront’s Maya 4, with its ghosting and motion trail features, was a great help. Most of you are familiar with unsettling feeling when viewing animation, in which you sense that something is wrong, but have a hard time determining exactly why the animation is not convincing. Seeing frames in onionskin allowed us to spot those mistakes much easier.

Additionally, the ability to switch between Inverse Kinematics (IK) and Forward Kinematics (FK) joint solutions was of great importance to the success of this animation. Motion is based on arcs, giving fluidity and grace to the character’s performance. It’s can be very difficult to get that kind of movement using IK solutions and an animator can spend a lot of time getting rid of the feeling that there is some magical force dragging the wrists and ankles of their character. Forward Kinematics at first seems to be much more time consuming but, in fact, it can make it easier to get the desired results.

While it may be difficult to see the work put forth by Bernd Angerer, Paul George, Keith Huggins, and myself, we believe that our efforts to animate the dying ringwraiths and horses helped add to the believability and success of this sequence.



Figure 1. Image of the Ringwraith Character Animation

# “Low Level” Intelligence for “Low Level” Character Animation

Damian Isla

Bruce Blumberg

MIT Media Lab, 20 Ames St.

Cambridge, MA 02139

{naimad, bruce}@media.mit.edu

## Abstract

It has been suggested that investing animated characters with low-level cognitive models can allow a rich set of low-level behavior to be produced automatically. For example, a character that has the cognitive capability of Object Persistence can intelligently direct its gaze over a scene and even respond emotionally to certain world events. This level of cognitive modeling allows for complete behavioral control by a human controller or a script, if that degree of control is necessary for the application.

**CR Categories:** I.2.0 (Artificial Intelligence): General – cognitive simulation

**Keywords:** AI and Behavior, Animation Techniques

## 1 “Low Level” Animation and Intelligence

We use the term “low level” to designate animation such as eye-motion, gaze control and facial expression – character animation, in other words, that has less to do with what a character is doing and more to do with how the character is feeling, the character’s personality, etc. Part of the irony of the term is that while the motion is physically minimal, its impact can be subtle and very powerful. It is also, in some cases, tedious or impossible to script beforehand, for example in interactive computer games where events must be intelligently recognized and reacted to on the fly.

An AI solution to the above might seem natural, since much of the low-level animation described is significant precisely because it is indicative of some kind of emotional or knowledge-state internal to the character. If a character frowns and continually glances towards a door, we might infer that it is because the character is anxious about someone soon coming through it. Much of the previous work in behavioral modeling has focused on the problem of action selection. This work, however, shows that cognitive models could be used on occasions where complete character autonomy is undesirable (where the activity of the character is scriptable beforehand, or is under a player’s control).

## 3 Object Persistence

As an example of a cognitive capability that can produce much interesting behavior, we will discuss the phenomenon of object persistence. If a character tracking an object sees the object pass behind an occluding obstacle, an unsophisticated character will simply say “I no longer know where the target is”. A slightly more sophisticated character might say “the target is at the location I last observed it.” A character with a full-blown model of object persistence would say “the target is behind the obstacle (and furthermore, based on the velocity it had when I last observed it, I expect it to come out again in x seconds).”

## 4 Implementation

A model of object persistence can be implemented as follows: the character’s environment is divided into a grid of discrete locations called a “Probabilistic Occupancy Map” (POM). This map is a two-dimensional discrete probability distribution representing the possible location of a single target object. When an observation of the target object is made, the grid element the target is in is given

a probability of 1. For every time-step of the simulation in which the object is not observed, each grid element diffuses some portion of the probability it contains to its immediate neighbors. Thus the longer the target is unobserved, the “fuzzier” the character’s idea of the target’s location. However, in maintaining this distribution, probability is always culled from locations that are currently wholly visible to the character (based on a limited field of view, and the inability to see through occluding obstacles). This ensures that the character’s ideas of where the target is remain “realistic”. Each time-step the distribution is also renormalized. Thus a location unlikely on one time-step might be much more likely the next, if the most likely location has been rejected (perhaps the character, in searching for the target, has gone behind the occluder only to find that the target is not there. In that case, the much less likely-seeming possibility that the target might have reversed directions and hidden in another location should be considered much more seriously).

## 5 Example Results

The above scheme was implemented in an animated character named Duncan, a virtual sheepdog. Duncan was instructed by a virtual shepherd to keep track of a sheep (both shepherd and sheep being controlled interactively) as it moved about a simple environment with two occluding walls. Duncan could be instructed to “look at the sheep”, “go to the sheep” and “go to the shepherd”. Interesting new behaviors included:

- Emergent look-around: by instructing Duncan to look at the sheep after he had not observed it for some time, Duncan used his POM to intelligently direct his gaze toward the sheep’s most likely location. When the sheep was especially elusive, Duncan would glance back and forth over the scene in a natural and evocative way.
- Emergent search: when instructed to “go to the sheep” (assuming the sheep was hidden behind one of the occluders) Duncan would go check out the most likely-seeming hiding place. Not finding it there, he would systematically search the environment until the sheep was found (or until he became so frustrated that he gave up).
- Secondary Emotions: by taking into account the likelihood of the location the sheep was eventually found in, Duncan could be made to display certain secondary emotions such as surprise (to have found the sheep in a place it was not expected to be) and confusion (at NOT finding the sheep in a place it was expected to be).

While only a very preliminary example of “low level” intelligence for “low level” animation, we believe that this level of cognitive modeling could prove extremely fruitful for crafting the next generation of intelligent (and intelligent-LOOKING) characters.

## References

ISLA, D. AND BLUMBERG, B. 2002. Object Persistence for Synthetic Creatures. In *Proceedings of the International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, Bologna, Italy.

# Magical face: Integrated Tool for Muscle Based Facial Animation

Tatsuo Yotsukura  
Seikei University/SEGA Corporation  
3-3-1 Kichijo-ji Kitamachi, Musashino-shi  
Tokyo, JAPAN  
yotsu@ee.seikei.ac.jp

Mitsunori Takahashi  
Shigeo Morishima  
Seikei University  
Kazunori Nakamura  
Hirokazu Kudoh  
SEGA Corporation

## Introduction

In recent years, tremendous advances have been achieved in the 3D computer graphics used in the entertainment industry, and in the semiconductor technologies used to fabricate graphics chips and CPUs. However, although good reproduction of facial expressions is possible through 3D CG, the creation of realistic expressions and mouth motion is not a simple task.

In fact, producing such effects requires a tremendous amount of time and effort on the part of the CG creator because there has been no tool that allowed automation in the initial creative process or the reuse of high-quality facial expressions. As a result, the reproduction of facial expressions is an expensive process and the quality of the finished product depends on the skill level of the creators.

To solve these problems, we developed Magical Face, an integrated tool that can be used for facial animation by users of any level. This tool consists of three sub-tools - a facial expression editing tool that uses an anatomical muscle model, - a face model fitting tool, and lip-sync animation tool that is based on the user's voice. These operate through Maya plug-ins (Alias-Wavefront Inc.).

## Facial Expression Editing Tool

This tool is based on a facial muscle model [Waters and Frisbie 1995; Lee et al. 1995] composed of facial tissue and the simulated muscles of a geometric model, which is arranged according to the anatomical configuration of the face. The muscular tissue is modeled as an aggregate of springs. In this model, forces affect a facial tissue element through the contraction of each muscle spring, so the combination of the contracting forces of various muscles produces a specific facial expression.

The control panel of the facial expression editing tool has slider bars for each muscle used to create a facial expression. When these bars are moved, each muscle detects the change in the applied force and the model is deformed. To realize motion that leads to a more natural facial expression, we have to adjust the target model and muscles by using a geometric model and move each slider bar to determine the appropriate muscle parameters for various expressions and mouth movements. The basic expression of Anger is shown in Figure 1.

## Face Model Fitting Tool

When a specific expression of a facial model has been developed with the facial expression editing tool, we normally would still have to set up the geometry and muscle parameters again if we wanted to reuse that expression. However, the face model fitting tool allows us to avoid this process. Therefore, we propose the Multistage Ray-Cast Fitting plug-in software. The user selects a face model whose muscle parameters were defined using the tool (the source model), selects a different number of polygons and a different polygon geometry (the target model), and then executes the program that approximates the source model

to the target model. Since the geometry of the deformed source model already has the muscle parameters, the specific expression and mouth shape of the target model can be easily formed. As well as a realistic human facial model, this model can use an animal's or a cartoon character's face as a target model. Figure 2 shows the adjustment result for several models.

## Lip-sync Animation Tool

When determining the animation of the mouth shape to apply to the face model made by the face model fitting tool, the model would normally need to speak while synchronized with the voice of an actor or actress. Also, if this model has to closely resemble the real face, the number of mouth shapes that must be prepared will equal the number of voice visemes. To do this is very time consuming for the creator who must pay close attention to the voice actor/actress whose voice must be synchronized with the mouth animation. The lip-sync animation tool makes this task easier by analyzing a sound and extracting the phoneme and phoneme length. This phoneme is then matched with the created mouth shape. To create a natural looking mouth shape animation, the phoneme is interpolated using the phoneme duration and the muscle's equation of motion. Currently, this tool can estimate each Japanese vowel. Therefore, the lip-sync animation tool significantly reduces the difficulty of applying a user's voice which previously required a series of complicated and time-consuming tasks.

## References

Waters, K., and Frisbie, J. 1995. A coordinate muscle model for speech animation. Proc. Graphics INterface '95, pp.163-170.  
Lee, Y., Terzopoulos, D., and Waters, K. 1995. Realistic modeling for facial animation. Proceedings of SIGGRAPH '95, pp. 55-62.

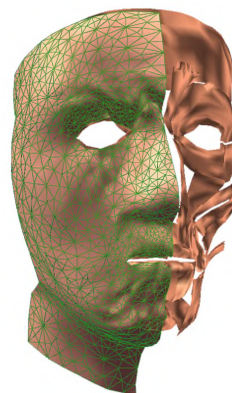
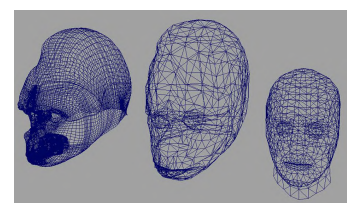
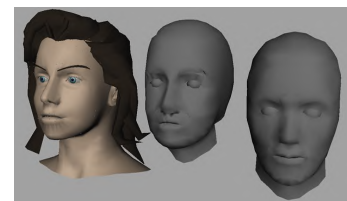


Fig.1 Facial expression of Anger



a) Gorilla model



b) Woman model

Fig.2 Face Model Fitting Tool

Left: Target model  
Middle: Reconstructed model  
Right: Source model



# MARE : Multiuser Augmented Reality Environment on table setup

Raphaël Grasset, Jean-Dominique Gascuel \*  
iMAGIS-GRAVIR, INRIA Rhône-Alpes  
655 Avenue de l'Europe, 38330 Montbonnot, France  
<http://www-imagis.imag.fr>

## 1 Introduction

Using augmented reality (AR) in collaborative situations is appealing: it combines the use of natural metaphors of communication (gesture, voice, expression), with the power of virtual ones (simulation, animation, persistent data). But few 3D AR collaborative systems are devoted to keep human's ability (like grasping, writing). The motivation of this research is to mix together virtual reality techniques ([Schmalstieg et al. 2000]) and computer human interaction techniques ([Fjeld et al. 2002]), so to have the best of both worlds.



Two users with the system.

A user's vision.

## 2 A testbed for interaction techniques

We present a completely open architecture (XML config files, OpenGL Performer), dedicated to the evaluation of different new interactive techniques. The system reproduces the configuration of a round-the-table meeting: it allows collaborators to keep contact with the real environment, while interacting with virtual objects. We decompose the table space in two parts: first, the personal area where the user put his private real objects (pen, PDA, laptop, cup) and virtual ones (private menu). The second part is the shared area, communication and common interactive space. In this space the user can:

- interact with virtual and real objects,
- add, remove virtual objects,
- register real objects to the system can account for them.

We present our system and some innovative solutions for these three tasks.

We introduce the concept of **least intrusive devices**: each user is equipped with devices, that try to keep all liberty of movement for the task in which the user is involved : navigation and interaction with virtual objects, perception and interaction with the real world. We use stereo optical see-through Head Mounted Display attached with magnetic or optical trackers. We also developed a **"mixed-pen" metaphor**: a pen you can write with, but also tracked so that you can use it to interact with virtual objects. We use a **wacom tablet** to realize annotation, gesture recognition, or manipulate electronic documents. We also define intuitive methods, in the spirit of "plug'n play", for an easy and **simple calibration** of the different tools and objects allowing to rapidly start and stop the sessions. Navigation is natural, thanks to **tracked HMD**. But we also

use the **Dive-in-the-model** metaphor to switch between AR and VR modes.

We developed different techniques of interaction based on what we call **"low haptic feedback"** (contact with the table, and preservation of gravity). The procedure to move an object is inspired by two metaphors: **dragging** an object on surface or **lifting** an object above an obstacle. We also developed a hand-attached device that can replace the "mixed-pen". Real and virtual objects are moved using the forefinger with the same drag/lift metaphor.

For adding virtual objects, we propose three methods: hierarchical virtual menus (in the private area of the user), gesture recognition (unistroke approach) and the catalogue metaphor (similar to VOMAR[Kato et al. 2000]).

**Adding real objects** can be used to mix them with virtual mockup, to use them for game surroundings, or as "tangible user interface". We propose two methods: if we have a CAO model, we ask the user to **superpose** a virtual representation of the real object to find its position. If not, we ask the user to **digitalize** the real object by inserting some simple geometric primitives (cube, cylinder, sphere, plane).

For **multiuser sessions**, we have realized an **access control manager**, to control personal view information (each user can have a different representation of the same object), to set permission to move or modify objects etc.

We integrated basic editing tools (move, rotate, scale, copy, delete) with a replay event manager and an help, information and annotation service (audio, video, image and wacom input). Our system can be easily used in **any domain of application** (users just have to customize interface and domain objects). We show some example applications: scientific application, engineering, architecture, urban planning, and game.

## 3 Future Works

To encompass the limited accuracy of magnetic trackers, a new high precision tracking system (*optotrak*) will be used and even less intrusive tracking methods, based on computer vision, will be tested. In complement, video can be used to develop new kinds of HMD (hybrid video/optical see-through) and to enhance the user vision (creating shadow/lighting effects).

In a similar manner to virtual reality, usability of augmented reality in collaborative application will be studied, during a cooperation with CHI specialists. Better interaction techniques are needed: "tangible user interface" and new 3D interaction wireless devices seem promising.

## References

- FIELD, M., LAUCHE, K., BICHEL, M., VOORHORST, F., KRUEGER, H., AND RAUTERBERG, M. 2002. Physical and virtual tools: Activity theory applied to the design of groupware. *Computer Supported Cooperative Work (CSCW) 11*, 1.
- KATO, H., BILLINGHURST, M., POUPYREV, I., IMAMOTO, K., AND TACHIBANA, K. 2000. Virtual object manipulation on a table-top ar environment. In *Proceedings of ISAR 2000*.
- SCHMALSTIEG, D., FUHRMANN, A., AND HESINA, G. 2000. Bridging multiple user interface dimensions with augmented reality. In *Proceedings of ISAR 2000*.

\*e-mail: {Raphael.Grasset/Jean-Dominique.Gascuel}@imag.fr

# MasterMotion: Full Body Wireless Virtual Reality for Tai Chi\*

Philo Tan Chua      Rebecca Crivella      Bo Daly      Ning Hu      Russ Schaaf      David Ventura  
Todd Camill      Jessica Hodgins      Randy Pausch  
Entertainment Technology Center, Carnegie Mellon University



Figure 1: Top-left to bottom-right: A student in the trainer. The virtual world. Center of gravity display. Coloring the limb with the largest error

We present a wireless virtual reality system and MasterMotion, the full-body training application built with it. We combine real-time full-body optical motion capture with wireless audio/video broadcast, belt-worn electronics, and a lightweight head-mounted display (HMD), to provide a wide-area, untethered virtual environment system that allows exploration of new application areas.

The application we have chosen is Tai Chi training. Tai Chi is composed of slow, fluid, motions, making on-line feedback and correction possible. This is a challenging training application because the sequence of movements, the *form*, is complicated and the standards for performance are exacting. Novices struggle to memorize the sequence of movements but more experienced students spend in excess of 20 years refining their movements.

## 1 System

Our training environment uses a Vicon real-time optical motion capture system from Oxford Metrics. We track students in a volume of 4 meters by 5 meters by 2.3 meters high and produce a 42 degree of freedom body representation. Students wear a wireless HMD through which they see a rendered virtual environment containing animated representations of both the student and the teacher. The HMD is extremely light (3.4 oz), which is important for a task involving body posture. We transmit an NTSC signal to the HMD via an off-the-shelf wireless video transmitter and receiver. The student wears a fabric hip belt that holds the batteries and the receiver. The total weight of the system is 2.5 pounds.

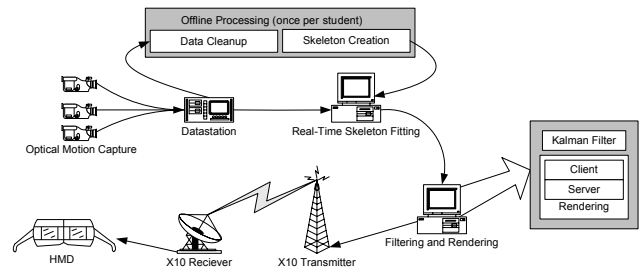


Figure 2: Data pipeline and system diagram

## 2 Feedback

It is critical that students learning Tai Chi are able to compare their motions to those of the teacher. To aid in this comparison, we implemented a number of feedback techniques.

The simplest feedback technique is a pure visual comparison between the student's body and the teacher's. To afford this comparison, we place multiple copies of the student and the teacher throughout the space. We also render the master as a wireframe or stick figure, so that the student can treat the teacher as a skin to fit inside, or as the student's skeleton.

For display of student error, we have implemented two techniques. The first, body coloring, uses an error metric to determine which of the student's limbs is most in error. If that error is above a configurable threshold, we color the limb red, giving students an immediate indication that they should correct the limb. The second technique is a center of gravity measurement. In Tai Chi balance and awareness of the center of gravity are very important. To visualize this, we render a heads-up display to the viewport showing the centers of gravity for the teacher and student relative to their feet. The center of gravity is approximated by projecting the pelvis position onto the line connecting the left and right ankle.

The final learning aid we have implemented is scalable audio annotation. We recorded a Tai Chi teacher giving commentary on the form. We then classified the audio segments based on their relevance to beginner, intermediate, and advanced student. At run-time, the system determines which audio clips to play based on the student's expertise level as well as the playback rate, allowing more instruction when the form is played slower.

## 3 Other Applications

Other physical training tasks that could benefit from this technology are dance, heavy lifting, and other repetitive motion training. Another approach may be necessary for motions which are faster than the ones we are working with. For quick motions it may be necessary to give students feedback in an offline manner, such as plotting the trajectory of the student's last swing of a golf club alongside that of a golf professional after the student has swung.

\*<http://www.etc.cmu.edu/projects/mastermotion>

# Maximum Entropy Light Source Placement

S. Gumhold  
University of Tübingen

## 1 Introduction

Several methods have been proposed to help the user to position lighting sources for a given view of a 3d-scene. In the first kind of methods the user defines a desired illumination through highlight and shadow locations for which the lighting system optimizes the light positions. The more automatic approaches are based on perceptual image metrics. [Marks et al. 1997] define an image metric that measures how different two images are perceived. A collection of maximally different images is presented to the user for selection. [Shacked and Lischinski 2001] define a perceptual based image quality metric composed of six contributing terms, for which the user has to specify weights before the system searches for a locally optimal light source placement.

## 2 Information Theoretic Quality Metric

This work evolved in the context of scientific visualization, where the primary goal is to disclose as much information about the dataset as possible. In the exploration stage of the visualization process, the user has only a rough idea of the dataset and cannot specify a target illumination. Our objective was the design of a fully automatic light source placer. Following the paradigms of visualization we chose to maximize the information content added to the final image by the illumination.

The information content of a given illumination is called the *illumination entropy*. For its computation the range of possible brightness values is partitioned into  $k$  equal sized subintervals  $b_{i=1\dots k}$ . The illumination entropy  $H = -\sum_i p_i \cdot \log p_i$  is computed from the probabilities  $p_i$ , that a brightness value of the given illumination falls into the interval  $b_i$ . Our light placing algorithm searches for a global maximum of the illumination entropy in the space of light positions – for example  $S^2$  in the case of a directional light source.

For the placement of more than one light we introduce the *conditional illumination entropy* of a light source. It measures the information content added by the light source under the condition that all other light sources are already turned on and is computed from the corresponding conditional probabilities. Using the sum over the conditional entropies as quality metric ensures independent contributions of the different light sources.

## 3 Adaptive Global Maximization

For any number of light sources, the quality metric is calculated from probabilities, which are computed by counting for each bin  $b_i$  the number of illuminated pixels whose brightness value is in  $b_i$ . Thus the quality metric is not differentiable with respect to the light source locations, what prohibits gradient-based optimization methods. Instead, a global optimization method was designed for the weaker assumption of Lipschitz continuity implying that the gradient is bound by a constant  $L$ .

The basic idea of the global optimization procedure is to subdivide the search space of light positions into a coarse grid and evaluate the quality metric on the grid vertices. Iteration by iteration the maximum quality metric and the Lipschitz constant  $L$  are estimated from the current sampling. All cells are subdivided, on

which the measured maximum value can be exceeded if the quality metric grows with the maximum gradient  $L$ . After one level of subdivision the quality metric is evaluated on the newly introduced vertices. The level wise iteration terminates when the maximum reachable value of the quality metric is within a user specified tolerance of the measured maximum.

## 4 Implementation and Results

Although we restricted the lighting computations to the Phong illumination model allowing the use of graphics hardware, the most time consuming step in evaluation of the quality measure is the rendering of the illumination. To reduce the computational complexity from scene to image complexity the illumination parameters of each covered pixel were extracted in three rendering passes allowing 40 illuminations per second in software on a PIII, 1GHz at a resolution of  $400^2$  pixels. Further acceleration was achieved by building a hierarchy in image space and by vector quantization of the illumination parameters, such that the optimal light locations could be found for a single directional light source in half a second.

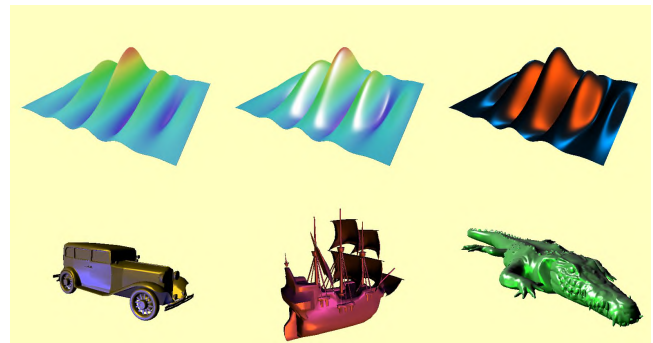


Figure 1: top row: bad illumination vs. maximum entropy placement of one and two directional light sources; bottom row: 3d-models illuminated with two directional light sources

Figure 1 shows examples of placing one and two light sources. The adaptive global optimization for the single directional light source took 0.45 seconds and for the other examples with two directional light sources 2.34, 1.98, 2.12 and 2.31 seconds respectively. Maximum entropy light source placement also produces appealing results for typical 3d-models as the bottom row of Figure 1 illustrates.

## References

- MARKS, J., ANDALMAN, B., BEARDSLEY, P. A., FREEMAN, W., GIBSON, S., HODGINS, J., KANG, T., MIRTICH, B., PFISTER, H., RUML, W., RYALL, K., SEIMS, J., AND SHIEBER, S. 1997. Design galleries: A general approach to setting parameters for computer graphics and animation. In *SIGGRAPH 97*, 389–400.
- SHACKED, R., AND LISCHINSKI, D. 2001. Automatic lighting design using a perceptual quality metric. In *EUROGRAPHICS 2001*, 215–226.



# MOCAP Game Reserve: A Study of Puppetry and Motion Capture

Charlotte Belland  
ACCAD, The Ohio State University  
belland.2@osu.edu

This animation sketch presents a study of human-driven character animation with motion capture as presented as part of the 2002 SIGGRAPH Course *Motion Capture: Pipeline, Applications, and Use*.

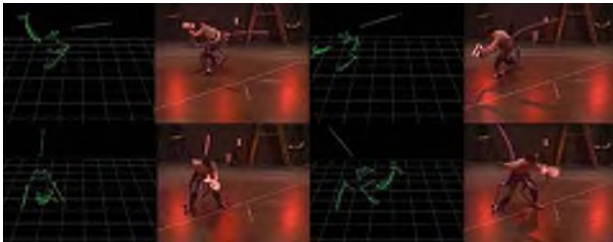
## 1.1 Goal

The goal is to establish a procedure of rapid puppet armature construction that encourages a modest amount of spontaneity. The methodology towards creature design is based from the actor's pelvis outward onto the armature. This interdisciplinary procedure communicates to individuals whose professions may not include computer animation but whose medium is grounded in a visual style (in this case, actors).

## 2.1 Movement and Marker Methodology

To keep the foundation of movement rooted on the actor, the pelvis or midsection of every character is marked directly onto the actor. These are the only four markers to have direct contact with the human body. The pelvis then serves as the primary source of motion for every character. The armature spreads from the pelvis outwards. All markers independent of the pelvis four must be located on extensions.

The hierarchy of motion places the pelvis of the actor as the most dominant primary motion. Other primary motions relate to direct manipulation of the actor onto the armature (such as armature attached to the feet or attached to a hand). Secondary motion is defined as any part of the armature that is not in direct contact with the actor (such as a tail or neck).



## 2.2 Visualization Tools

A construction material was needed for physical puppet design that could endure multiple motion capture takes, be flexible enough to prevent bodily harm but rigid enough to show solid form. After experimenting with a variety of building supplies, an interesting tool emerged: the Pool Noodle. These flexible pool toys, made of closed cell polyethylene foam, provided the perfect raw building material since they were easy to cut (a mat knife and a coping saw), were lightweight, could be held together with duct tape and were easily accessible from any major store that carried children's toys.

This building tool and the realities of using humans as the foundation of motion thereby affect character design. The

proportions and stance of the resulting characters must be sized to fit the actors. This provides an interesting design challenge when it comes to the decisions of where and how joints will rotate. In the case of this study, the armature for the legs was produced so that the anatomy of the leg mimicked a bird leg. The armature, due to its flexible nature, added an additional range of secondary motion that was inherent only to this building tool. The motion of the character's leg was enhanced with an additional wiggle to its step.

## 2.3 Application of Methodology and Tools

The first creature to emerge in the MOCAP Game Reserve was a bird-like character. The first four markers were placed directly on the actor's pelvis. Next, Pool Noodles were used as extensions to create the legs. The Pool Noodles were cut on angles at the "joint" so that there was a limited range of rotation with no hyper extended joints. Six separate sections of Pool Noodle were strung together to create the neck. The base of the neck was fastened to the belt of the actor and the top of the neck was fastened to a T-joint head. The tail was an intact Pool Noodle held behind the actor.

## 3.1 It Lives

Multiple motion takes were recorded, illustrating the bird in a variety of motions; walking, running, noticing, searching and dancing. A selection of data was cleaned with Vicon's WorkStation. Only gaps were filled and the rest of the data was left to exist in its original format. The data was then processed in House of Moves DIVA™ to create a skeleton that was imported into Alias|Wavefront MAYA™ with the House of Moves Dominatrix plug-in. Since this was an experiment in motion, a model composed of virtual Pool Noodles was created to honor the versatility of the armature tool.

While there was no formal protocol for review, informal viewings from faculty, staff and students from Art, Design, Computer Science and Theatre departments generated excitement. The motion was instantly recognized as a character independent of its creation.

The energy created by this first creature has led to its offspring's addition into a proposed Theatre production for the following academic year.





# Modal Analysis for Real-Time Viscoelastic Deformation

Chen Shen Kris K. Hauser Christine M. Gatchalian James F. O'Brien  
University of California, Berkeley  
{csh,khauser,tine,job}@eecs.berkeley.edu

## Introduction

This technical sketch describes how a standard analysis technique known as modal decomposition can be used for real-time modeling of viscoelastic deformation. While most prior work on interactive deformation has relied on geometrically simple models and advantageously selected material parameters to achieve interactive speeds, the approach described here has two qualities that we believe should be required of a real-time deformation method: the simulation cost is decoupled from both the model's geometric complexity and from stiffness of the material's parameters. Additionally, the simulation may be advanced at arbitrarily large time-steps without introducing objectionable errors such as artificial damping.

## Background

Modal decomposition is a commonly used tool for analyzing physical systems. A thorough review of the technique may be found in [Maia and Silva, 1998], but we will describe it very briefly here.

A physical system that has been discretized using a finite element, finite differencing, or other similar method can be expressed in the following general form:

$$\mathbf{K}\mathbf{d} + \mathbf{C}\dot{\mathbf{d}} + \mathbf{M}\ddot{\mathbf{d}} = \mathbf{f} \quad (1)$$

where  $\mathbf{d}$  is the vector of node displacements, an over-dot indicates a derivative with respect to time, and  $\mathbf{K}$ ,  $\mathbf{C}$ , are respectively known as the system's stiffness, damping, and mass matrices. In general the system matrices are not constant, but if the expected deformations are small then the system may be linearized. This linearization is the main limitation of the technique we describe. It means that we cannot use it for materials such as cloth, but we can use it for solid rubbery materials.

Modal decomposition treats (1) as a generalized eigenvalue problem and transforms it into a diagonalized system such as

$$\mathbf{\Lambda}\mathbf{z} + (\alpha_1\mathbf{\Lambda} + \alpha_2\mathbf{I})\dot{\mathbf{z}} + \ddot{\mathbf{z}} = \mathbf{g} \quad (2)$$

where  $\mathbf{\Lambda}$  is a diagonal matrix of eigenvalues,  $\mathbf{z}$  is the displacement vector expressed in the eigen-basis, and  $\alpha_1$  and  $\alpha_2$  are the Raleigh coefficients.

This idea has been used previously in the graphics community for modeling deformation by [Pentland and Williams, 1989], however they did not actually use a modal decomposition, they only approximated it with global linear and quadratic basis functions. When done properly, (2) is *not* an approximation of (1), it is *exactly* (1), but expressed in a much more efficient basis. Furthermore since (2) is decoupled, analytical solutions may be used for each of the modes so that numerical time integration is no longer necessary. Instead the object's behavior may be computed by convolving the analytic solutions with any externally applied forces.

## Real-Time Modeling

To model an object using modal analysis first require having a suitable discrete model of the object. We are using the finite element formulation described by [O'Brien and Hodgins, 1999] to compute the system matrices. The eigen-system is then computed using [Wu and Simon, 1999]. These two steps may require a few hours, but they are pre-computation that only needs to be done once for a given object.

At run-time, our system uses the modal decomposition for computing the response of an object to user interaction. Only the modes that are over-damped, critically-damped, or under-damped with a frequency below half the frame rate are used. Under-damped modes with a frequency greater than half the frame rate would only create temporal aliasing.

The behavior of each retained mode is modeled with a pair of complex oscillators. Interaction forces from the user are projected onto the eigen-basis and used to excite the oscillators. Because the oscillators do not require numerical time integration, we can run the simulation at arbitrarily large time-steps. The computation cost of the simulation is determined by the number of retained modes, not the geometric complexity of the model.

We have implemented this method. A 4000 vertex model can be modeled in real-time on a 500 MHz laptop PC. The bulk of the computation time is consumed by the graphics functions, not the simulation.

## References

- MAIA, N. M. M., AND SILVA, J. M. M., Eds. 1998. *Theoretical and Experimental Modal Analysis*. Research Studies Press, Hertfordshire, England.
- O'BRIEN, J. F., AND HODGINS, J. K. 1999. Graphical modeling and animation of brittle fracture. In *Proceedings of SIGGRAPH 99*, 137–146.
- PENTLAND, A., AND WILLIAMS, J. 1989. Good vibrations: Modal dynamics for graphics and animation. In *Proceedings of SIGGRAPH 89*, 215–222.
- WU, K., AND SIMON, H. 1999. TRLAN user guide. Tech. Rep. LBNL-42953, Lawrence Berkeley National Laboratory.

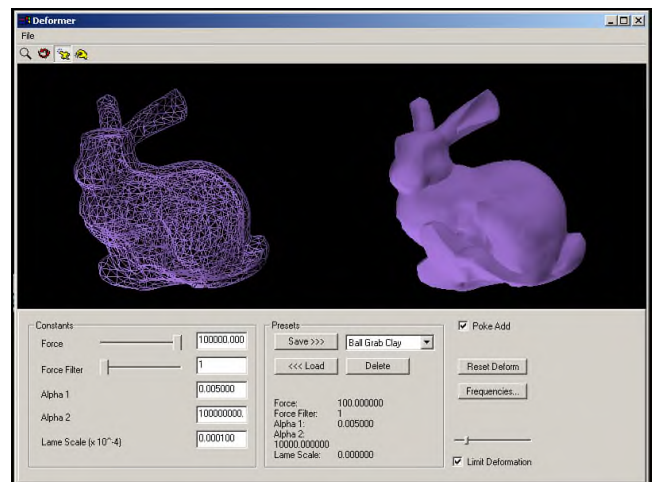


Figure 1: A screen shot from an interactive deformation application.

# Modeling the Accumulation of Wind-Driven Snow

Bryan E. Feldman   James F. O'Brien  
University of California, Berkeley  
bfeldman@me.vehicle.berkeley.edu, job@acm.org

## 1 Introduction

This technical sketch presents a method for modeling the appearance of snow drifts formed by the accumulation of wind-blown snow near buildings and other obstacles. Our method combines previous work on snow accumulation [Fearing] with techniques for incompressible fluid flows [Fedkiw et al.]. By computing the three-dimensional flow of air in the volume around the obstacles our method is able to model how the snow is convected, deposited, and lifted by the wind. The results demonstrate realistic snow accumulation patterns with deep windward and leeward drifts, furrows, and low accumulation in wind shadowed areas. (See figure.)

## 2 Methods

The first step of our method is to compute a steady flow field for the wind around the obstacles. As the snow accumulates, this flow field will be treated as constant until a significant amount of snow has accumulated. As large amounts of snow accumulate we periodically recompute the flow to reflect resulting changes in wind patterns.

The space around the obstacles is discretized into a uniform Eulerian grid. Cells within the obstacles are marked as impermeable. Cells directly above upward facing horizontal surfaces are marked as landing sites.

We compute the flow fields using the Navier-Stokes equations for three-dimensional inviscid, incompressible flow. The cells on the boundary of the three-dimensional grid are initialized with the desired average wind velocity. A sequence of alternating mass-conserving and momentum-conserving steps are computed using a semi-Lagrangian integration scheme. (See [Fedkiw et al.]) The sequence terminates when the flow field has converged so that the maximum magnitude change in wind velocities from one time step to the next is below a threshold.

Once the flow field has been computed, the next step requires computing how snow moves through the flow and accumulates on the landing sites. Snow is injected into the scene by setting the border cells to contain an amount of snow equal to  $\rho \hat{n} \cdot \mathbf{v}$  where  $\rho$  is the density of snowfall,  $\hat{n}$  is the inward facing normal of the cell's exterior face(s), and  $\mathbf{v}$  is the average velocity of the falling snow away from the obstacles. If this quantity is negative, then snow is not injected at that border cell.

The snow inside the grid is convected by moving it in a direction which is the sum of the local fluid velocity and a downward gravitational term. A landing cell removes a portion of its snow and stores it as resting on the corresponding horizontal surface. If the wind velocity in a landing cell is large, then a portion of snow will be removed from the horizontal surface and returned to the flow field.

The amount of snow at each landing site is compared to its neighbors to ensure that some critical slope is not exceeded. If the critical slope is exceeded between a site and its neighbor, the neighbor is checked to see if it can receive the snow. Impermeable cells cannot receive snow. If a neighbor is free to accept the snow, it receives a quantity proportional to the amount the critical slope is exceeded plus some small amount to ensure convergence. Neighboring open sites have the avalanched snow injected into the flow, while neighbor landing sites add the avalanched snow to the corresponding surface. The process repeats until all landing sites are stable.

The snow convection process proceeds until a suitable amount of snow has accumulated in the scene. If large drifts that might alter the wind flow form, then the flow field may be recomputed.

Once we have a final distribution of accumulated snow, the accumulation on the horizontal surfaces is used to generate a height field. The height field is rendered as a Catmull-Clark subdivision surface.

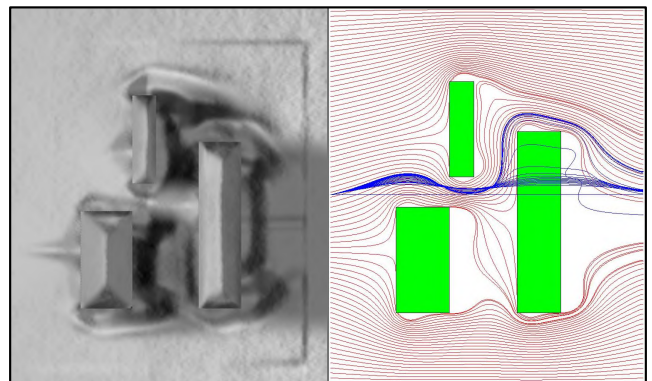
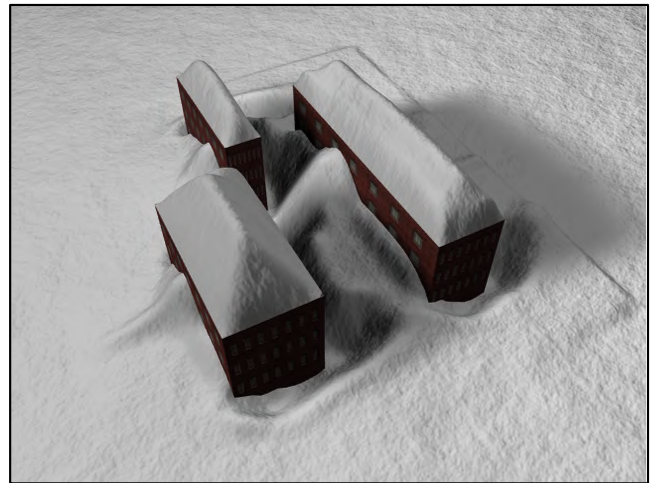
## 3 Results

We have tested our method on the simple scene shown in the figure. While this scene is much simpler than those shown in [Fearing], the use of a flow field to compute the snow's motion allows us to capture effects that were not producible by the previous method. Some of these effects include long windward drifts, furrows at the sides of obstacles, and leeward gullies and drifts. The top image shows a perspective view of a large amount of snow accumulating on and around 3 m tall obstacles. The lower images show a top view of the same scene and a top view of a set of streamlines through the flow field. Approximately three hours were needed to generate these results with matlab on a laptop PC.

We would like to thank Christine M. Gatchalian for assisting with rendering these images.

## References

- FEARING, P. 2000. Computer modelling of fallen snow. In *Proceedings of ACM SIGGRAPH 2000*, 37–46.
- FEDKIW, R., STAM, J., AND JENSEN, H. W. 2001. Visual simulation of smoke. In *Proceedings of ACM SIGGRAPH 2001*, 15–22.



# Modified Marching Octahedra for Optimal Regular Meshes

Hamish Carr \*

Thomas Theußl †

Torsten Möller ‡

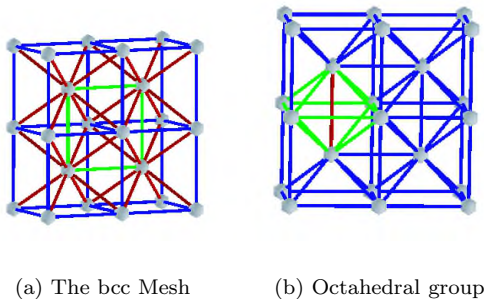


Figure 1: The body-centred cubic mesh

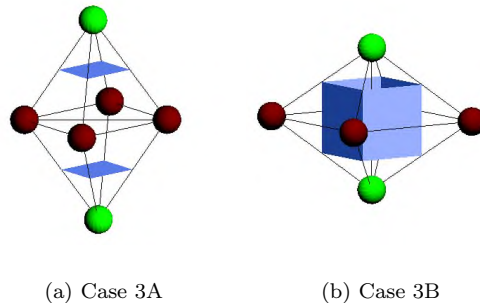


Figure 2: Case 3 of Modified Marching Octahedra

## 1 Mesh Construction

Theußl et al. [Theußl et al. 2001] showed that volumetric data sampled on a body-centred cubic (BCC) lattice is nearly 30% more efficient than data sampled on a cubic lattice, and produced volume renderings using splatting. We extend this work to generate isosurfaces based on the BCC lattice, and also on the hexagonal-close packed (HCP) grid. This sketch presents a modified version of marching octahedra that simplifies the BCC mesh to an octahedral mesh to reduce the number of triangles generated for the isosurface.

The *marching cubes* technique [Lorenson and Cline 1987; Wyvill et al. 1986] uses a cubic cells: the Delaunay complex of the samples. The Delaunay complex for HCP uses tetrahedra (*tets*) and octahedra (*octs*). The Delaunay complex for BCC (Fig. 1(a)) has two superimposed cubic meshes, shown in blue (*primary lattice*) and green (*secondary lattice*), with diagonals between the two meshes, shown in red.

## 2 (Modified) Marching Octahedra

Tetrahedral cells can be dealt with by marching tets [Bloomenthal 1988]. We define six cases for marching octs using the same methodology as for marching cubes and marching tets. Only one case is ambiguous (Fig. 2). The faces of an oct are triangular, not square, so the cracks of marching cubes [Dürst 1988] do not occur, and we can choose either case without causing cracks in the surface. Although tets use fewer triangles than cubes for isosurfaces, there are many more tets in the BCC grid than cubes in the cubic grid. We address this by grouping tets into octs.

In the BCC lattice, octs composed of groups of four tets share each edge of the secondary lattice, as shown in

\*Department of Computer Science, University of British Columbia, Vancouver, BC, Canada. [hcarr@cs.ubc.ca](mailto:hcarr@cs.ubc.ca)

†Institute of Computer Graphics and Algorithms, Vienna University of Technology [theussl@cg.tuwien.ac.at](mailto:theussl@cg.tuwien.ac.at)

‡Department of Computing Science, Simon Fraser University, Vancouver, BC, Canada. [torsten@cs.sfu.ca](mailto:torsten@cs.sfu.ca)

Fig. 1(b). We refer to the shared edge as the *spine* of the oct. If the two vertices “above” the isosurface are along the spine of the oct, marching tets renders the surfaces shown in Fig. 2(b). If not, marching tets renders the surfaces shown in Fig. 2(a). We use the location of the spine to choose case 3A or 3B, guaranteeing that the surface generated is topologically equivalent to that generated by marching tets. By substituting octs for tets, we attempt to reduce the number of triangles rendered for an isosurface on a BCC lattice.

## 3 Expected Triangle Counts

We estimate the number of triangles generated by assuming that  $O(N^{2/3})$  cells intersect each isosurface, ignoring constants. We adjust  $N$  for BCC and HCP to reflect the 29.3% saving in samples noted above, and multiply by an estimate of the number of triangles generated per cell - 2.85 for cubes, 1.25 for tets, 4.07 for octs (modified). This gives the following estimates for similar quality sampling:  $2.85N^{2/3}$  (cubes),  $3.42N^{2/3}$  (HCP),  $3.75N^{2/3}$  (BCC),  $4.23N^{2/3}$  (BCC with octs). We are evaluating triangle counts experimentally: preliminary results are similar to these estimates.

## References

- BLOOMENTHAL, J. 1988. Polygonization of implicit surfaces. *Computer Aided Geometric Design*, 341–355.
- DÜRST, M. 1988. Letters: Additional Reference to “Marching Cubes”. *Computer Graphics* 22, 4, 65–74.
- LORENSON, W. E., AND CLINE, H. E. 1987. Marching Cubes: A High Resolution 3D Surface Construction Algorithm. *Computer Graphics* 21, 4, 163–169.
- THEUSSL, T., MÖLLER, T., AND GRÖLLER, M. E. 2001. Optimal Regular Volume Sampling. In *Proceedings of Visualization 01*, IEEE, 91–98.
- WYVILL, G., MCPHEETERS, C., AND WYVILL, B. 1986. Data Structure for Soft Objects. *Visual Computer* 2, 227–234.



# Motion capture done dirt cheap

Stefan Gustavson (stegu@itn.liu.se)  
 Linköping University, Campus Norrköping  
 SE-601 74 Norrköping, Sweden

Modern real time motion capture systems are complex, large scale installations. They focus on accuracy, reliability, ease of use for technical staff and convenience for performers, but unfortunately they are also prohibitively expensive for many applications where motion capture might otherwise prove useful. Low budget work and experiments, VR applications and university research could all benefit from a low cost alternative for motion capture. Such applications could also tolerate a somewhat lower quality, and possibly even some slight inconvenience for users and performers.

We have designed a very simple and cheap motion capture system with good accuracy, based on physical angle sensors (conductive plastic potentiometers) mounted on a skeleton rig. Our prototype skeleton rig is puppet sized, but alternatively, using the same hardware, a human size exoskeleton could be designed for full body motion capture of an actor.

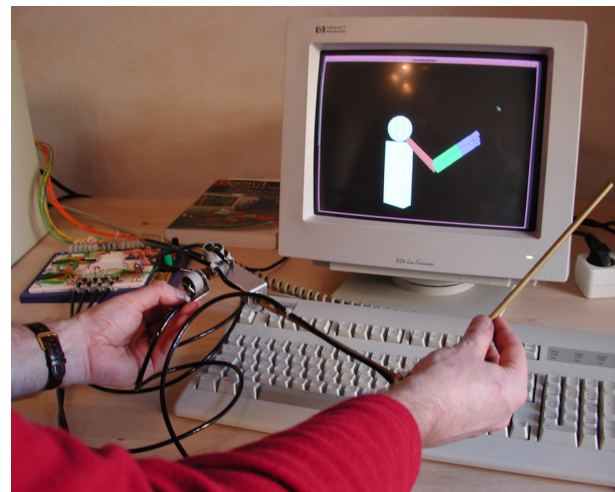
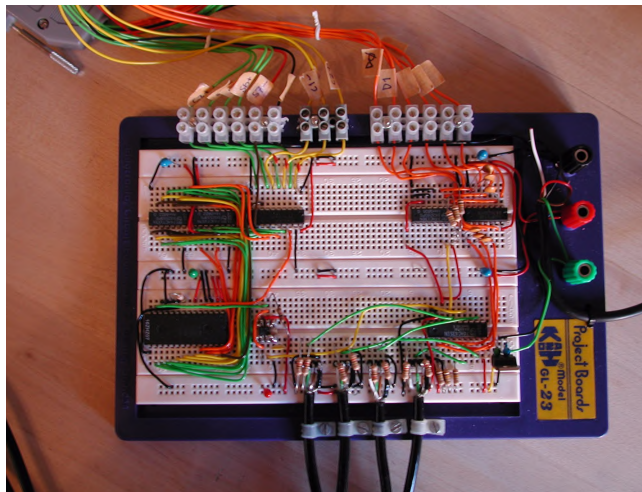
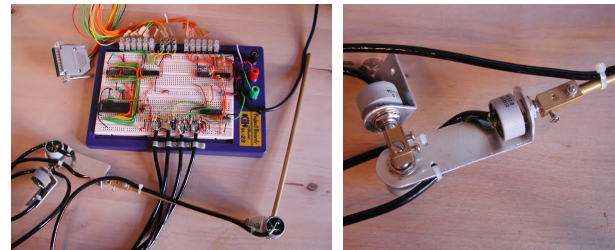
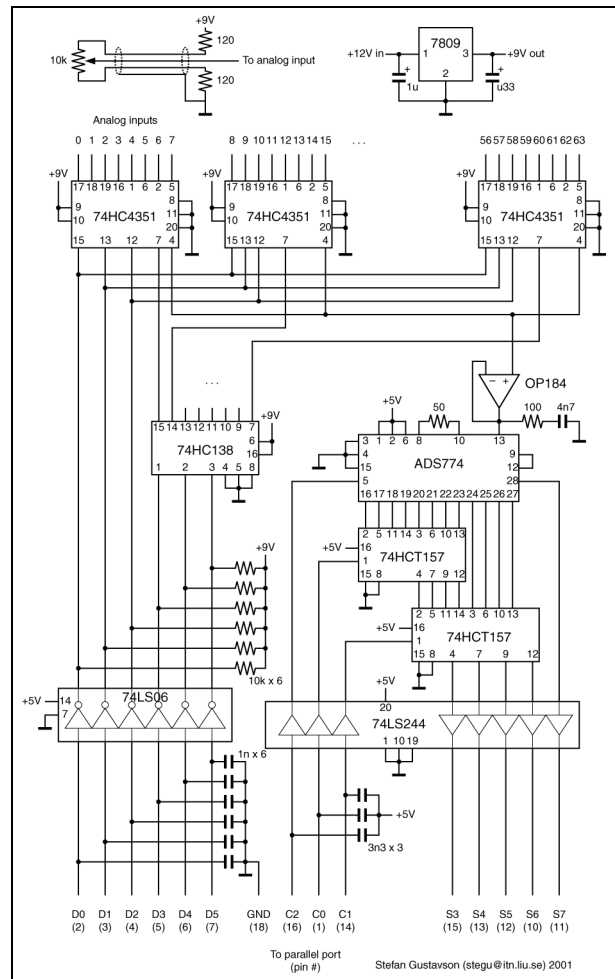
The system connects to the parallel port of a standard PC and is currently capable of sensing up to 64 analog channels of 12 bit angle data at a sample rate of 60,000 channel samples per second. A 64 channel setup yields a maximum sample rate of around 1 kHz. Fewer channels allow for proportionally higher sample rates.

The total hardware cost of our prototype system with eight angle sensors was less than US \$300. Additional angle sensors are a mere \$10 each. The system was assembled and put into operation in one week by one person. Thus, motion capture does not necessarily have to be expensive.

Some experience with electronics and basic mechanics design is required to build the device, but only hobby level tools are needed, and the design is quite simple.

Our current system is a prototype and a technology demo, not a turn-key solution to any real world problem, but it provides us with a reasonably good and very inexpensive platform for further experiments. We will demonstrate a simple application to animate an OpenGL model in real time using motion capture input.

A full description of the prototype with schematics, parts list and source code for the software is available on:  
<http://www.itn.liu.se/~stegu/monkey/>





# Motion-Based Shape Illustration

Eric B. Lum

Aleksander Stoppel  
University of California, Davis  
{lume,stoppel,ma}@cs.ucdavis.edu

Kwan-Liu Ma

## 1 Introduction

Motion provides strong visual cues for the perception of shape and depth, as demonstrated by cognitive scientists and visual artists. We present a novel visualization technique that uses particle systems to add supplemental motion cues. Based on a set of rules following perceptual and physical principles, particles flowing over the surface of an object not only bring out, but also attract attention to essential shape information of the object that might not be readily visible with conventional rendering. Replacing still images with animations in this fashion, we show with both surface and volumetric models that the resulting visualizations effectively enhance the perception of three-dimensional shape and structure.

This work is motivated by the observation that the flow of fast moving water over a rock, a dynamic flame from an open fire, or even a flock of birds exhibit motion that gives the perception of shape. Our technique is built on the inspirations we received from kinetic art, the studies done in cognitive science, the ideas of particle systems, and the previous work using texture to convey the shape of overlapping transparent surfaces.

## 2 Motion Strategies

To generate geometrically meaningful motion, the particles are governed by a set of rules. The overall goal is to create rules resulting in particles that indicate shape by smoothly flowing over an object, with locally consistent directions, and a density distribution that does not let particles "clump" together in regions of little interest.

Since we would like to better illustrate an object's shape, rules are imposed to constrain the motion of particles to be along a surface. The motion of particles along an object's surface can help improve shape perception, over time presenting the viewer with a set of vectors (trajectories) that run parallel to a surface.

The principal curvature directions indicate the direction of minimum and maximum curvature on a surface. We use principal curvature directions to create particles that "follow the shape" of a surface. Particle velocities are adjusted so the particles flow in a direction that favors one of the principal curvature directions chosen by the user.

Since particles moving in opposite directions can provide confusing motion cues, we use a set of rules that move particles in directions consistent with their neighbors. One method we use to give the particles more consistent orientations is to give the particles flock-like behavior. Flocks exhibit motion that are fluid, with each member still exhibiting individual behavior. Thus flocking can be used to add local uniformity to particle motion while still allowing particles to have motion shaped by outside forces like principal curvature direction. The other method for giving particles a consistent orientation is to simply define a "preferred" direction the particles must move, for example having all particle have a preference toward downward motion.

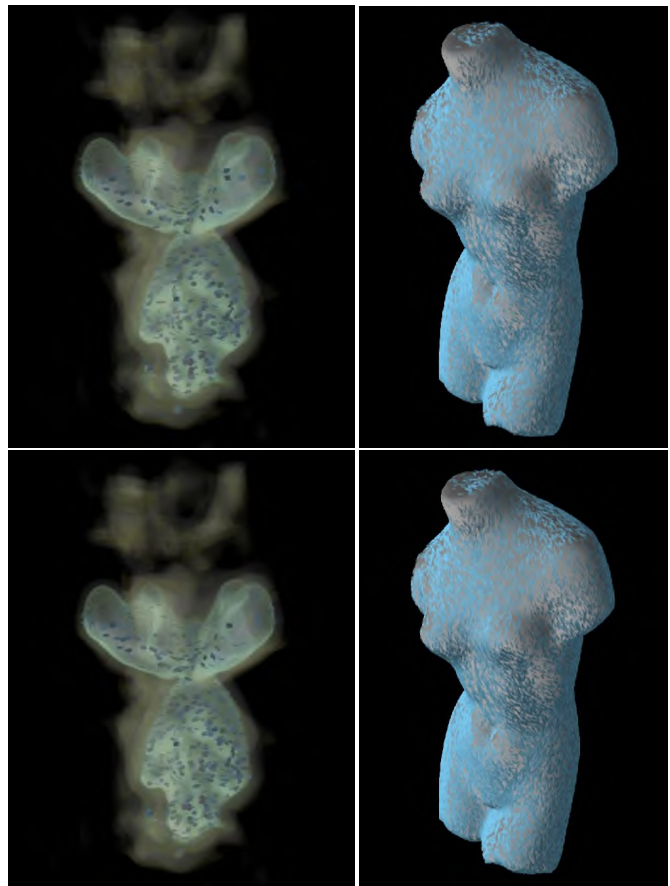
Since the number of particles has a direct influence on rendering time, it is desirable to have a set of rules that efficiently uses a limited budget of particles. By using a set of rules based on magnetic repulsion, more uniform particle densities can be achieved. Particles are modeled as having a magnetic charge of the same sign, and

are repelled from their neighbors with a force inversely proportional to the square of the distance of the neighbors.

## 3 Demonstrations

To demonstrate this motion-based technique, several animation sequences were generated and can be downloaded from <http://www.cs.ucdavis.edu/~ma/KinVis>. Two frames from a sequence for the visualization of a mouse brain PET volumetric dataset are shown in the left column below. The particles help to illustrate one of the function levels while volume rendering gives their motion context. On the right column, two frames from another sequence which shows the application of rules to better illustrate a surface model of Venus. The particles flow down the Venus model in a manner similar to water. The downward tendency adds consistency to the motion.

The proposed technique can create more perceptually effective visualizations by adding visually rich motion cues. While more work will be required, our current results are encouraging, demonstrating that it is feasible and desirable to capitalize on motion cues for the purpose of enhancing perception of 3D shape and spatial relationships.



# Music-Driven Motion Editing

Marc Cardle  
mpc33@cam.ac.uk

Stephen Brooks  
sb329@cam.ac.uk

Loic Barthe  
lb282@cam.ac.uk

Mo Hassan  
mfh20@cam.ac.uk

Peter Robinson  
pr@cam.ac.uk

Computer Laboratory  
University of Cambridge

## Introduction

Sounds are generally associated with motion events in the real world, and as a result there is an intimate connection between the two. Hence, producing effective animations requires the correct association of sound and motion, which remains an essential, yet difficult, task. But unlike prior, application-specific systems such as [Lytle90; Singer97], we address this problem with a general framework for synchronizing motion curves to perceptual cues extracted from the music. The user is able to modify existing motions rather than needing to incorporate unadapted musical motions into animations. An additional fundamental feature of our system is the use of music analysis techniques on complementary MIDI and analog audio representations of the same soundtrack.

## System Overview

In our system, MIDI and analogue audio features visually influence existing motion curves such as keyframe animation and motion capture. Hence, the system consists of three distinct modules: MIDI, Audio and Motion editing. Along with an initial motion sequence, the output of the two music analysis modules is used by the motion editing module to alter the final motion (Figure 1). Selected music analysis parameters modulate the chosen motion editing methods. These motion alterations are layered to encompass multiple aspects of the music simultaneously. The user-guided mapping process continues until the animator is content with the resulting animation. Our system thereby adapts to user preference as opposed to previous fully automated systems in which the user is required to adapt to the resulting output.

## Music Analysis

The system provides a comprehensive selection of music analysis techniques and permits iterative testing of diverse combinations of music and motion editing parameters. An iterative approach is necessary for two reasons: firstly, the animator may not know precisely what they are trying to achieve and require some leeway for experimentation; secondly it may not be obvious which parameter mappings are required to achieve particular effect once it has been decided on. To this end, a range of perceptually significant features need to be extracted from the music. This is partly a process of reducing the overwhelming acoustical information and focusing on aspects that may be desirable to translate visually.

Since the conversion from raw MIDI data into a synthesized soundtrack has a crucial influence upon the final perception of the music, we include both MIDI and analogue audio-based features. Examples of such features are pitch induction, chord recognition, segmentation and pattern matching in the MIDI domain; spectral energy, peaks and flux and envelopes are examples of audio-based features. Variations of these features are mapped to variations in the motion.

## Motion Editing

When tightly controlled by music analysis, current motion editing techniques offer wide scope for motion control using a variety of standard signal processing tools. For instance, frequency spectrum rescaling, linear and non-linear filtering, can all be made to affect the emotional content of the animation. Filter banks can be used to divide a motion signal into a number of components, which can be manipulated independently according to music parameters and then reassembled to create a new motion signal. Time and displacement warping, multi-resolution editing techniques and multi-target blending provide distinct yet mutually compatible alterations to the motion. Combinations of these operations would be prohibitively complex to carry out manually.

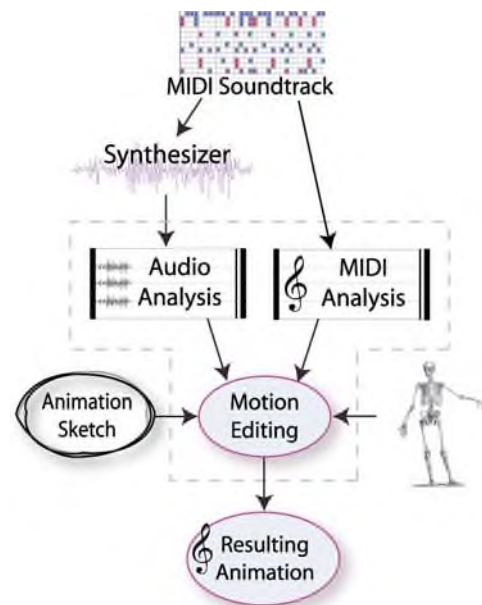


Figure 1: System Overview

## Conclusion & Future Work

Early results indicate our system streamlines the workflow of integrating music and animation. Additional work to be completed includes a number of motion editing techniques, as well as a flexible GUI which will allow us to rapidly and extensively test multiple combination of music and motion parameters. Lastly, human motion-specific transforms will be investigated.

## References

- LYTLE, W. 1990. Driving Computer Graphics Animation From A Musical Score. *Scientific Excellence in Supercomputing: IBM 1990 Papers*.
- SINGER, E., AND ROWE, R., 1997. Two highly integrated real-time music and graphics performance systems. *ICMC*.

## Now and then, here and there - Industrilandskapet

Konstantin Economou, Mark Ollila, Martin Etherton Friberg, and Anders Ynnerman  
Norrköping Visualization and Interaction Studio  
University of Linköping  
Norrköping Campus, Sweden  
kosec@ituf.itn.liu, marol@itn.liu.se, martin.etherton@ituf.itn.liu, andyn@itn.liu.se

As cultural heritages go, or rather stand firm, the Norrköping “industrial landscape” (industrilandskapet in Swedish) has been the object of cultural regeneration over the last two decades. Once, one of Sweden’s largest city center industrial sites, major supplier of textile and paper it has now become the nexus of cultural activity. Museums, a concert hall and, the university, have now made the scenery different. In turning the landscape over we might think that old industrial life has succumbed to the changes in industrialisation making Norrköping move from manufacture to knowledge industry. In museums, this era of “old” is depicted in nostalgic ways – as “lost” cultural heritage rather than as living imagery. And, where did manufacturing go? Is it the case that industry has changed or has it just moved.

By connecting images of early 1900 industrial life with contemporary images from the textile industry in India we explore the connection of now and then; and of here and there. In an exhibition project that will extend the museum space to comprise the university Virtual Reality theatre, the internet and an outdoors location in the industrial landscape itself, we use these four locations to create an extended and linked space of experience and exhibition, thus also making a statement in contemporary museum theory and practice. The project is thus linked to contemporary Swedish and international debate on the role of museums in society and the development of subject matter and representation in track with contemporary societal issues and debates. A statement is made and position taken on issues of the city’s structural and historical changes.

The museum (Norrköping city museum) is in this project turned inside out and connected to three more real and virtual spaces. The project is debating traditional cultural content moving from just collecting representative objects to representing vivid history related to the narrative and construction of contemporary interpretations of (in this case industrial) social change.

This is made possible by the use of new media technology not only in the sense that stories hitherto untold, images and film combined with augmented reality gives new possibilities of time and space travel, but also because their new, non-linear and non-space bound narrative possibilities can be used to break the traditional viewer-, or visitor-, expectations that traditional museum narratives and displays have sometimes stuck with. In this sense the creation of the “visit” to this project will take into account the creation of the visitors experience guiding a future visitor, thus making any person interacting with the exhibit a possible producer for the “next fellow audience”.

The technology that is been used in this project is based on many new research findings published in the literature. Image Based Rendering and Modelling technology is used in generating realistic looking models from the photographs assembled over time. This is used for offline rendering, as well in combination with real time techniques such as Relief Texturing which will be displayed in the Virtual Reality Theatre. Augmented reality is explored in producing real experiences in the environment itself, and thus experimenting



Figure: From top moving clockwise. 1. Snapshot of Industrilandskapet today. Park on left of image use to be a factory in 1935. Augmented Reality exhibition will be used at that location, showing the old factory as well as zooming into the snapshots of India. 2. Snapshot inside a factory in 1928 (textile industry and dying factory). 3. Photograph of a women from Norrköping washing clothing. 4. Image-based rendering used for representing the region. 5. Image of people in the region of Norrköping in the late 19th and early 20th century. 6. Images from today in India (textile industry). 7. Image of person dying textiles in India in present day.

More information can be obtained from the website  
<http://www.nvis.itn.liu.se/industrilandskapet>



# Occidio

Timothy Nohe  
University of Maryland Baltimore County  
FA111, 1000 Hilltop Circle,  
Baltimore, MD 21250 USA  
[nohe@umbc.edu](mailto:nohe@umbc.edu)

## Abstract

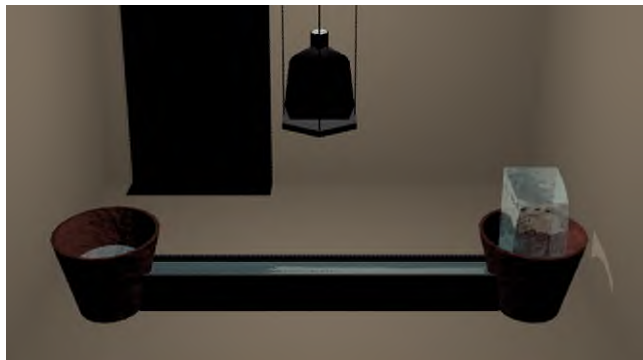
This sketch presents an artwork entitled *Occidio*, which has been produced to interpret the alarming phenomena of global warming in the form of a computer-controlled sound and DVD installation. *Occidio* interprets NASA scientific visualizations through the interplay of video projections, event-triggered synthetic sound generated by optical theremins, and metaphoric sculptural forms. Visitors to *Occidio* shall "see," *and*, "hear" interpretations of global warming.

## Sound

The sound of global warming is performed by optical theremins stimulated by projected visualizations. Optical theremins are a very simple type of sound synthesizer, which produce a variable pitched sound wave when light stimulates a photoreceptor. The theremin's photoreceptors are to be placed directly in the light path of the video projection. Hue and value changes would modulate the audio tones to "perform" different sounds when triggered by changes in global warming projection. A computer applying real time granular synthesis signal processing would transform the audio signal into chromatic and rhythmic phrases, separating the sound into a sonic environment. This environment would be reproduced through a 4-channel amplification system, with sound reproduced through four discreet speakers. I consider the abstract yet elegiac sound resulting from this work to be a form of "music," if one interprets the term conceptually, and, as an alternative and very direct means of interpreting scientific data.

## Form and Installation

The dominant visual element in the installation interior is a symmetrical sculptural form (plate, below) consisting of two vessels interlinked by an open channel. The structure is designed to contain the runoff of a solid 100-lb. block of ice. The ice would be placed in one vessel at the opening of the installation; over time the ice would melt, filling the channel and each vessel equally. Earth's accelerating snow and glacier melts are alluded to through this process as the ice block slowly transforms from a solid to a liquid state.



The second visual element consists of global warming data projected onto the ceiling of the installation space (plate, below). This global warming data consists of public domain visualizations gathered from NASA data sets. The edited DVD loop would include visualizations related to changes in atmospheric chemistry, the polar "ozone holes," biosphere degradation, changing sea temperatures, glacier melts, etc.

In the proposed installation I would employ a video projector; a looping DVD player; optical theremin sound emitting devices; a computer mediating the theremin sound forms with real time granular synthesis; a sound mixer; a four channel audio amplifier and four speakers.



## Synthesis

*Occidio* presents a very direct way to interpret and understand the fundamental changes taking place on Earth through the synthesis of visual, metaphoric and sonic elements. The goal of *Occidio* is therefore didactic, to introduce the abstract but essential findings on global warming to the public through alternative means of visual, poetic and sonic interpretation. To reach that audience *Occidio* may be installed in both traditional and non-traditional venues: colleges, airports, galleries and museums, empty storefronts in urban spaces, etc. The work is designed to affect audiences directly in installation interactions, and through tertiary experiences via the reception and reproduction of the work in media such as the press, radio, television and Internet.

See also:

Occidio <http://research.umbc.edu/~nohe/OCCIDIO>

NASA Earth Observatory <http://earthobservatory.nasa.gov/>

NASA GCMD <http://gcmd.gsfc.nasa.gov/>

GOME <http://auc.dfd.dlr.de/GOME/index.html>



# Painterly Fire

Saty Raghavachary  
DreamWorks Feature Animation  
saty@anim.dreamworks.com

Fernando Benitez  
DreamWorks Feature Animation  
fbenitez@anim.dreamworks.com

## Abstract

This sketch describes how to construct a painterly 'wall' of fire, one which exhibits realistic motion while managing to maintain an artistic look.

**CR Categories:** I.3.5 [Computer Graphics]: Computational Geometry and Object Modeling—Splines; I.3.6 [Computer Graphics]: Methodology and Techniques—Graphics data structures and data types; I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Animation

**Keywords:** effects animation, rendering, natural phenomena

## 1 Introduction

For our third animated feature “Spirit - Stallion of the Cimmaron”, we needed to create a ravaging fire which threatens the life of Spirit and other animals in the forest. This sketch describes our technique used to achieve this objective.

## 2 Technique

We build upon 'Spryticle', a sprite system created inside Maya by our colleague Doug Ikeler, for use in “The Road to ElDorado”<sup>1</sup>. Spryticle predominantly has flat cards that always face the camera (the cards do allow for some bulge deformation). In contrast, our technique, called 'SpriteRibbons'<sup>2</sup> creates multi-segment flexible sprites analogous to streamers, and the segments are free to be oriented along arbitrary directions.

The gist of our process is as follows. The 'backbone' of each sprite is generated as a space curve by linking together instantaneous particle world-space locations across a series of frames. This backbone gets expanded out to a multi-segment sprite on which traditional hand-drawn texture cycles are mapped.

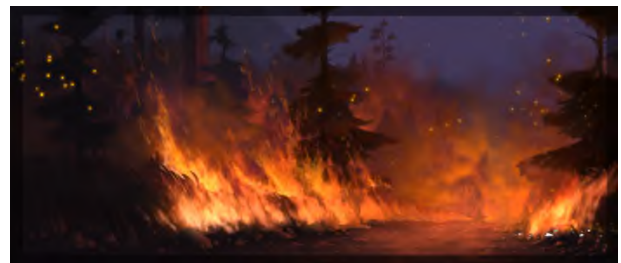
We begin by emitting particles in Maya near the ground plane. The particles are guided along upward paths by a Maya plugin called 'dwPathField'<sup>3</sup>. These particles have extra attributes such as

sprite-scale, sprite-normal, etc. to be used at the ribbon generation stage. The particle data is exported to disk in the form of Dynamation '.pda' files. Next a MEL script called 'SpriteRibbons.mel' is invoked to specify parameters for the sprite ribbon generation. These parameters include the length of the ribbons (in terms of particle frame numbers relative to the current frame), patch type (bilinear or bicubic) and whether or not at each frame, the sprite segments should each have their scale/normal/color/opacity or simply borrow the values of the central segment. A standalone program 'create-SpriteData' reads the .pda files and from the ribbon-generation information specified above, creates '.xpda' files (a custom extension of the Dynamation .pda format). The .xpda files contain normal and tangent information, making it easy for the next step (rendering) to generate sprite patch geometry.

The sprites themselves are generated in RenderMan at render-time, using a procedural modeling DSO<sup>4</sup>. The alternative to such just-in-time geometry creation would be to create them in Maya, with the attendant increase in scene complexity, scene file size and exported RIB file size. In MTOR, a 'RIBBox' is inserted with a call to the DSO (“RManSprites.so”). The DSO needs just the name of the .xpda file at each frame, and MTOR inserts this during RIB export. It reads the .xpda files and creates either bilinear or bicubic spline patches as specified by the user. Standard RenderMan surface shaders can be used to shade the sprites. We color the sprites using ramps, and use hand-painted flame shape texture cycles to matte out the colors. This gives a pleasing appearance to each sprite as it winds up its trajectory, changing flame color, shape and opacity as it does.

By varying the texture-mapped flame shapes, adjusting the length of the sprites, layering sprites with differing lengths and adjusting their widths, we are able to achieve a variety of looks for our fire.

The fire elements were used to color the background and also to warp it, both of which helped to integrate the elements into the scene. Further, smoke and embers were generated using the same particles which were used to create the sprites, and this added an extra touch of realism to the generated imagery.



<sup>1</sup>Doug Ikeler: *The Use of Spryticle in the Visual FX for “The Road to El Dorado”*, SIGGRAPH 2000 Conference Proceedings and Abstracts.

<sup>2</sup>Doug Cooper: internal communication.

<sup>3</sup>initially developed by Patrick Witting, extended and supported by Gigi Yates.

<sup>4</sup>[www.pixar.com](http://www.pixar.com), PhotoRealistic RenderMan Application Note #23.

# Painting with light

Kevin G Suffern\*

Faculty of Information Technology  
University of Technology, Sydney  
PO Box 123, Broadway  
NSW 2007  
AUSTRALIA

This sketch presents fractal art work created by ray tracing the specular highlights of point lights on the inside surface of a hollow sphere. The sphere has a mirror surface on the inside that contributes no colour to the images, but there is spread in the local specular reflection. The resulting images consist entirely of specular highlights and their reflections. I call this painting with light because, although recursive ray tracing is used, no objects are visible.

A bump map based on fractional Brownian motion [Musgrave 1998] is placed on the sphere to chaotically scatter the highlights and their reflections. For the results presented here, the bump map was applied three times by repeatedly feeding the perturbed normal back into the bump map. This was done because it produced the best results.

Figure 1 shows *Impact*, which is reminiscent of an asteroid impact. This was produced with three lights.



Figure 1: *Impact*, recursion depth = 10.

Light reflections inside a hollow sphere spread right across the images after a few bounces, particularly when a bump map is used. If we require high recursion depths, how can we restrict colours to the central parts of an image for artistic purposes? The answer is *depth lights*, where the light colour depends on the recursion depth. The red in the center of Figure 1 and the dark blue in the center of Figure 2 are examples.

Figures 1 and 2 use linear interpolation of the Perlin vector lattice noise [Perlin 1985] in the bump map, but I have also experimented with cubic interpolation. This is slow, but for a given set of fBm parameters, results in completely different images than linear interpolation, particularly when high recursion depths are used. Figure 3 was created with cubic interpolation.

There is a large parameter space to explore here. For example, the bump map parameters: number of octaves, lacunarity and gain; the viewing parameters; lights: number, locations, colours, and powers; and the optical properties of

the sphere. In addition, everything can have affine transformations applied, and there are other bump maps to explore.

Interesting animation effects can be created by varying the parameters, and applying affine transformations.

The software used is a custom ray tracer, which is essential for the non standard rendering techniques.

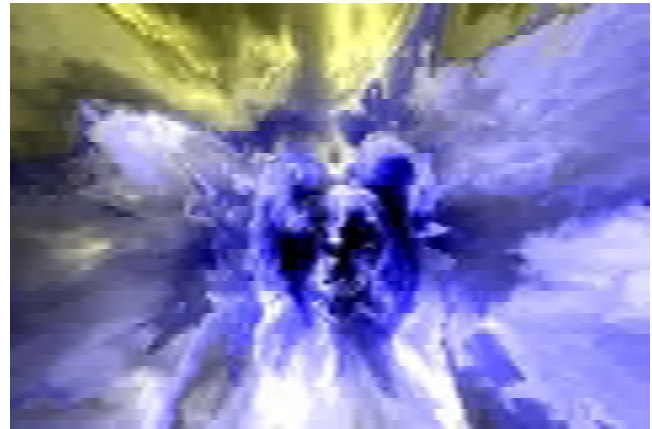


Figure 2: Use of depth lights, recursion depth = 12.

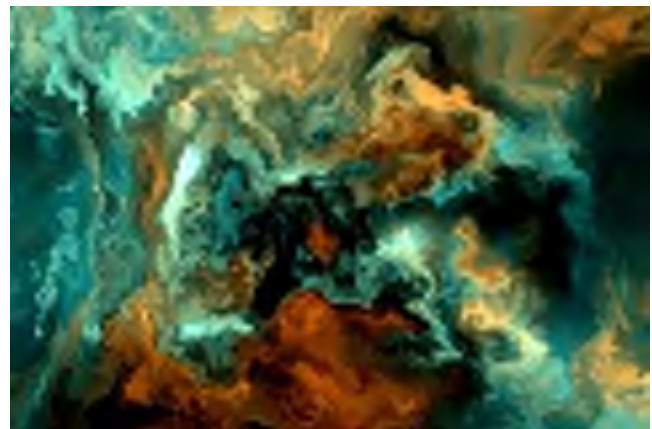


Figure 3: Cubic interpolation, recursion depth = 12.

## References

- MUSGRAVE F. K. 1998. A Brief Introduction to Fractals, in D. S. Ebert Ed., *Texturing and Modelling, A Procedural Approach*, Second Edition, AP Professional, Boston.
- PERLIN K. 1985. An Image Synthesizer, In *Proceedings of ACM SIGGRAPH 1985*, ACM Press / ACM SIGGRAPH, New York, B. A. Barsky Ed., Computer Graphics Proceedings, Annual Conference Series, ACM, 287-296.

\*e-mail: kevin@it.uts.edu.au

# Pastel-like Rendering Considering the Properties of Pigments and Support Medium

Kyoko Murakami  
Kyushu Institute of Design Graduate School  
kyoko@verygood.kyushu-id.ac.jp

Reiji Tsuruno  
Kyushu Institute of Design  
tsuruno@verygood.kyushu-id.ac.jp

## 1. Introduction

There has been any works specifically on computer generated 3D pastel rendering. Some rendering or imaging softwares can generate pastel drawing-like images recently, but most of them cannot modeling and rendering objects as 3D, and do not have enough power of expression in comparison with real-pastel drawings.

We will present a new NPR technique to generate pastel-like computer graphics image. In this technique, we reproduce pastel drawing-like texture by focusing especially on attributes of pastel pigment and the relationship between pigment and paper.

We treat objects as 3D model, it enables us to use traditional rendering preprocess or postprocess, such as transformation, projection, shading and animation.

## 2. Over view

First, we define 3D object formed by rectangle polygons. After that, we use common rendering steps such as transformation, projection, or hidden surface/line removal, we obtain surfaces or lines from which are not only original polygons but also a silhouettes and feature edges [Gooch and Gooch 2001].

Finally, assuming a 2D plane, we generate the pastel strokes with the pastel pigments, to obtain pastel drawing-like images.

## 3. Paper and Pigment

The texture of the papers that are mainly used in pastel drawing as support media is one of the most important elements. Therefore, we can get pastel-like effect by reflecting the paper unevenness in the whole picture as shown in Sousa et al.[1999]. Here, we change the paper unevenness into shading values of black-white and use the whiteness value as the pigment's adhesion. The pigments of pastel adhere on the high adhesion parts of the paper in response to the pressure on pastel.

The number of pigments at the pixel defines the pigment's transparency, because it is necessary to consider the color mixture of the paper and the pigment, in case the paper has a color itself.

## 4. Stroke

See figure 1. First of all, we define "basic line" of CG pastel stroke by splitting between origin  $s$  and endpoint  $e$  into a fixed length  $dl$  and placing the points (a). "Basic line" is given blur by staggering the position of the points randomly (b).

Next, to interpolate between the points in the "basic line", we generate "interpolate stroke". "Interpolate stroke" goes in the direction indicated by "stroke vector", that is defined from two adjacent points. Go to the direction that "stroke vector" indicates and put pigments on each place. Here, we define the amount of the pigments applied at once, that is the adhesiveness of the pigments, and determine the transparency of the pixels. If the length of the "interpolate stroke" is reached to  $dl$ , we end the generation (c).

After that, to create the stroke width, we apply the same technique for each side of the "interpolate stroke". In this process, we use the previously placed pigments as the origins and orthogonal oriented stroke vectors, which are given random blur. If the fixed width  $sw$ —that is blurred to make the stroke silhouette soft (e)—is reached, we

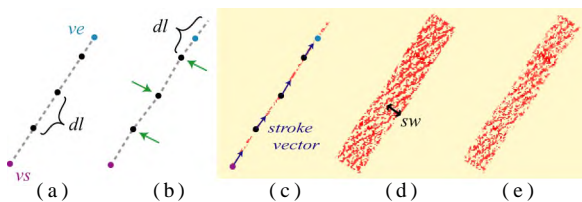


Fig 1. Generating Process of pastel strokes.

end the generation (d). When the strokes of different colors overlap on the paper, the colors are mixed in the pixel. Then, to make the color of the last stroke stronger than the others, its transparency is set to a lower level than the previous one.

## 5. Stroke techniques

To increase the representation power, we reproduced some common stroke techniques, as follows. The resulting images are in figure 2.

*Normal stroke*: Normal pastel stroke. Pastel pressure decreases from the origin to end point.

*Side stroke technique*: This is a wide pastel stroke. It is drawn using pastel body. To get better representational power, one of the edges of pastel is more pressured than the other side.

*Blending*: Blur the drawn strokes on the paper by finger and so on, to mix pigment color. It is also used simply to give a smooth look to the drawing.

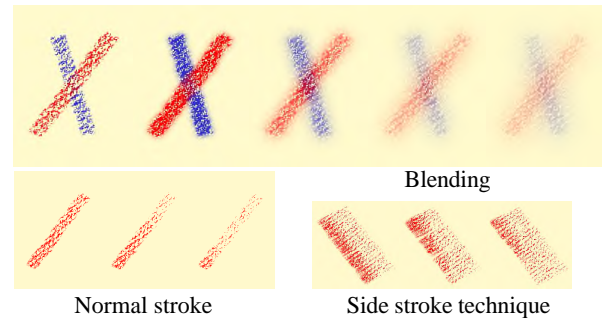


Fig 2. Stroke techniques

## 6. Results

Result images using this technique are shown in figure 3. These 3 boxes has same stroke parameters (stroke width, color, stroke technique) and different paper textures. And we applied *Blending* to the below images.

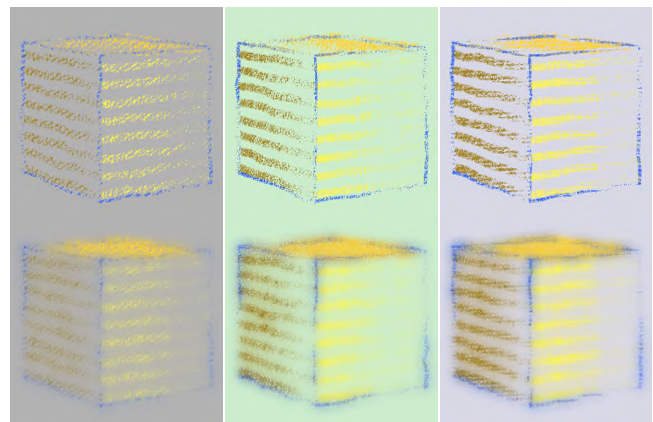


Fig 3. Result images

## References

- GOOCH, A., GOOCH, B. 2001. *Non-Photorealistic Rendering*. A.K.Peters.
- SOUSA, M. C., BUCHANAN, J. W. 1999. *Computer-Generated Graphite Pencil Rendering of 3D Polygonal Models*. Computer Graphics Forum 18(3), 195-208.
- TAKAGI, S., NAKAJIMA, M., AND FUJISHIRO, I. 2000. *Volumetric Modeling of Additional Techniques in Colored Pencil Drawing*. In Conference Abstracts and Applications of ACM SIGGRAPH2000 (Technical Sketches), 276.



# Perceptual Gaze Extent & Level Of Detail in VR: Looking Outside the Box

Hunter Murphy\*      Andrew T. Duchowski \*  
Department of Computer Science, Clemson University

A common assumption exploited in perceptual Virtual Reality studies is that eye movements made while immersed in VR generally do not deviate more than  $30^\circ$  (visual angle) from the head-centric view direction (e.g., see Barnes [1979]). In this sketch we report eye tracking evidence which generally supports this observation in the context of peripheral Level Of Detail management during a visual search task in VR. We present results from experiments based on the work of Watson et al. [1997] and discuss an extension to the peripheral degradation paradigm to include a dynamic eye-slaved high-resolution inset.

## EXPERIMENTAL TECHNIQUES

To investigate perceptual gaze extent in VR, we began by replicating the experiments of Watson et al., where the objective of the study was the evaluation of peripheral degradation for bandwidth reduction during a visual search task. Watson et al. found that a static high-resolution window of  $30^\circ$  in a low-resolution  $75.3^\circ \times 58.4^\circ$  visual field does not impede visual search performance (when compared to search in an undegraded environment). To quantify eye movement behavior under this condition, we use a binocular eye tracker mounted in an HMD (see Murphy and Duchowski [2001]). We then extend the static-window paradigm by synchronizing the high-resolution inset to the real-time location of the participant's gaze. Five trials were conducted for each of seven combinations of inset size and varying level of background degradation. This was repeated for both stationary and mobile insets resulting in 70 trials per subject. Eight subjects participated, varying in age from 22 to 53, with normal or corrected to normal vision.

## RESULTS AND ANALYSIS

In accordance with Watson's results, we found that a medium-sized high-resolution inset ( $< 30^\circ$ ) is required before search performance is impeded in a low-resolution scene. During this task, as shown in Figure 1, 70% of all fixations made by subjects fall within  $30^\circ$  (as indicated by the superimposed box). Our data shows that eye movements may extend as far as approximately  $43^\circ$ , consistent with Watson's assumption (based on the work of Barnes).

Since we found that windows smaller than  $30^\circ$  affect performance in a coarsely degraded field, we also investigated the use of an eye-slaved high-resolution inset. Figure 2 shows average trial search times. The first trial pair shows search times in a high-resolution field (no distinguishable inset or degradation). The next three trial pairs show no clear advantage of either the stationary or mobile insets. Surprisingly, two of the last three trial pairs show that search performance decreases when the high-resolution inset is synchronized with gaze direction.

## DISCUSSION

Our eye tracking work offers supporting evidence of perceptual gaze extent roughly limited to  $30^\circ$  during visual search in VR. This suggests that eye movements made outside the  $30^\circ$  box do not seem to affect search performance, i.e., subjects *look* but don't *see* outside the box. Furthermore, performance appears to suffer under extreme degradation conditions with a small, eye-slaved high-resolution inset. This may be due to either system lag, eye tracker inaccuracy, or imposition of interface control (gaze-controlled movement of the

inset) on a perceptual organ. We are continuing our analysis in order to disambiguate the latter point.

## REFERENCES

- BARNES, G. R. 1979. Vestibulo-Ocular Function During Co-ordinated Head and Eye Movements to Acquire Visual Targets. *Journal of Physiology*, 127-147.
- MURPHY, H., AND DUCHOWSKI, A. T. 2001. Gaze-Contingent Level Of Detail. In *EuroGraphics (Short Presentations)*, EuroGraphics.
- WATSON, B., WALKER, N., HODGES, L. F., AND WORDEN, A. 1997. Managing Level of Detail through Peripheral Degradation: Effects on Search Performance with a Head-Mounted Display. *ACM Transactions on Computer-Human Interaction 4*, 4 (December), 323-346.

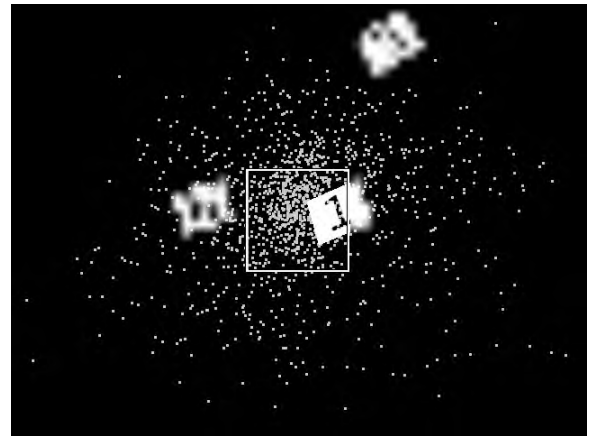


Figure 1: Fixations over an example search environment.

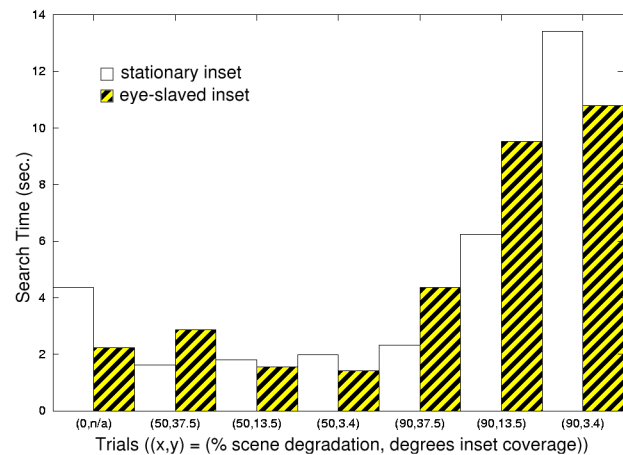


Figure 2: Average trial search times.

\*{hmurphy | andrewd}@vr.clemson.edu



# Perceptual Tone Mapping Operators for High Dynamic Range Scenes

Patrick Ledda and Alan Chalmers  
Department of Computer Science  
University of Bristol, UK  
ledda@cs.bris.ac.uk

Greg Ward  
Exponent  
USA  
gward@lmi.net

## 1. INTRODUCTION

A major goal of realistic image synthesis is to generate images that are both physically and perceptually indistinguishable from reality. One of the practical obstacles in reaching this goal is that the natural world exhibits a wide range of colors and intensities. The range of the luminances in the real world can vary from  $10^{-4}$ cd/m<sup>2</sup> (for starlight) to  $10^5$ cd/m<sup>2</sup> (for a daylight scene). Reproducing these luminances on a cathode-ray tube (CRT) display is currently not possible as the achievable intensities are about 100 cd/m<sup>2</sup> and the practical ratio between maximum and minimum pixel intensity is approximately 100:1. At the University of Bristol, we have constructed a High Dynamic Range (HDR) viewer that is capable of achieving a 10,000:1 contrast ratio. This sketch investigates, by means of psychophysical experiments, the benefits such a HDR device has to offer realistic computer graphics.

Realistic image synthesis has shown that it is possible, via tone mapping, to generate images similar to a real scene by careful mapping to a set of luminances that can be displayed on a CRT. Different tone mapping operators, for example [Ward et al 1997; Tumblin et al 1999], have been developed over the years, however none of them have managed to offer the substantial contrast reduction needed to display images without some loss of information.

## 2. PSYCHOPHYSICS

The goal of this study was to determine what the visible difference is between a standard display and a high dynamic range viewer. The HDR viewer we built for displaying stereographic, high dynamic range images is shown in figure 1. The viewer consists of a bright, uniform backlight joined to a set of LEEP ARV-1 optics allowing a 120 degree field of view on each eye. This gives a stereo view from a pair of fisheye transparencies mounted side-by-side, shown in figure 2. Due to the limitation in the film recording process it is not possible to achieve high dynamic range in a single process. Although the film is actually capable of producing about 1000:1 contrast, the useful range where the intensities steps are below the human perceived threshold are about 100:1. To solve this problem two transparencies layered on top of each other are used generating a dynamic range of approximately 10,000:1. The range is thus doubled in log terms since the densities are added together.

As well as rendered images with high dynamic ranges, photographs were also taken of scenes with a wide range of luminance levels. The high dynamic range was recovered using multiple exposures, [Debevec and Malik 1997; Ward 2001]. Both the rendered images and the photographs were tone mapped using the Ward et al [1997] and Tumblin et al [1999] operators. The resultant images were then showed to 40 subjects, 20 each on a standard gamma corrected monitor and 20 on the HDR viewer. The subjects were asked to identify carefully selected targets in the scenes, shown in figure 3 [Ayres 1996], and then answer a short questionnaire on the perceived quality of the images.

The results of the psychophysical experiments show the difference of performance of the tone mapping operators on both a CRT and the HDR viewer, indicating how future tone mapping operators should be developed in order to fully exploit the potential of high dynamic range viewers.

## REFERENCES

- AYRES T., Psychophysical Validation of Photographic Representations, ASME 1996.  
DEBEVEC, P.E., AND MALIK, J. 1997. Recovering High Dynamic Range Radiance Maps from Photographs. In *proceedings of ACM SIGGRAPH'97*, 369-379.  
TUMBLIN, J. AND HODGINGS, J.K., GUENTER, B. 1999. Two Methods for Display of High Contrast Images. *ACM Transactions on Graphics, Vol. 18, No. 1*, 56-94.  
WARD, G. HDRgen software, private communication, 2001.  
WARD, G. A Wide Field, High Dynamic Range, Stereographic Viewer. Exponent - Failure Analysis Assoc. (USA)  
WARD, G., RUSHMEIER, H. AND PIAKTO, C. 1997. A Visibility Matching Tone Reproduction Operator for High Dynamic Range Scenes. *IEEE Transactions on Visualization and Computer Graphics* 3, 4, 291-306.



Figure 1: HDR Viewer capable of 10,000:1 contrast ratio.

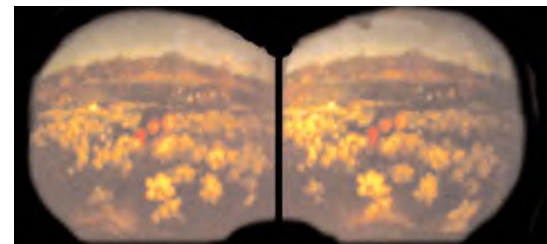


Figure 2: Fisheye transparencies for the HDR viewer.



Figure 3: Contrast chart [Ayres 1996].

# Performance driven computer graphics making *Odyssey*

Markus Manninen, Framestore CFC  
markus.manninen@framestore-cfc.com

## Introduction

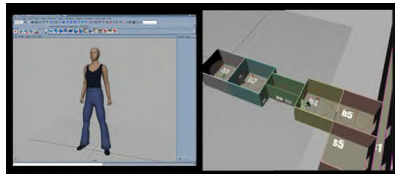
This animation sketch presents a production case study of how computer graphics work was designed to support and allow for director Jonathan Glazer's live action shooting style on "Odyssey", his latest internationally acclaimed commercial for Levi's. The film features a man and a woman in an ecstatic state of freedom to move. They achieve this by first escaping their drab interior surroundings, relentlessly running through a succession of walls. Once outside, they run vertically up two enormous trees and

upon reaching the very tip take a huge leap and launch themselves up towards the stars.



## CGI Previsualization

The effects team on *Odyssey* was involved from the early production meetings. The shooting style of the film would depend heavily on fluid camera moves tracking the fast moving performance of the actors. Therefore it was decided that CGI effects, featuring a computer generated forest and exploding walls, was the best solution. The CGI work began with previsualization to allow the set to be built based on the actors' performances. They were filmed running and hurdling, and based on this computer actors were built and animated. Another task of the previsualization was to enable Jonathan Glazer to work out shooting issues; speed, phasing, camera locations, movement, settings, and rough editing. The process allowed the director to press forward with the shoot the way he wanted to.

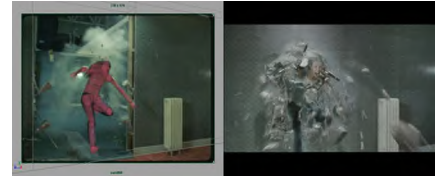


## Exploding CGI Walls

Previsualization determined the size and layout of the rooms based on the speed and a four-step hurdling of the actors. It helped the set designer who had to build a total of nine rooms and a corridor. Several methods were discussed to allow for the CGI to replace the walls of the rooms. The final solution was to have the live action room set built with large gaps between wall sections that would later be filled in CGI. First the walls were tracked back into plate in 2D (in *Discreet Inferno*). The 3D (in *Alias/Wavefront Maya*) used digital still images of the walls as textures rather than background plates, to allow for rendered motion blur. These wall passes were then blended together in final composite.

One of Jonathan Glazer's main concerns was how to get the performance of the actors across. In the wall sequence this meant the effects team having to solve the issue of the actors hitting a non-existent wall. The final solution was to have the actors running through a completely empty opening in the wall. A piece of tape was stuck across the opening indicating how high the actors should jump. This was slowly raised throughout the shoot without the actors' knowledge. Two dust cannons showered them from either side with dust and small debris to create real dust that was used in the final composite image together with additional CGI particle dust and debris.

The cannons achieved the appropriate reaction from the actors, as they were trying to avoid getting dust and debris in their faces and eyes. Also the dust stuck to their clothing, making them dirty as they moved through opening. This in turn allowed for the CGI wall pieces to integrate well with the live action. A procedural explosion system allowed for a first pass of animation. Then animators keyframed each individual piece of the explosion, with special consideration to the physical interaction of the actor and the wall pieces.



## CGI Forest

In the tree sequence the actors' performances meant enabling interaction between the actors and tree branches. The complete set of the forest was realized in CGI and the actors were performing in a blue screen set. The physical effects team added real tree branches in the path of the actors to allow for realistic interaction as they ran through the branches. When the CGI trees were added in post most of these branches were painted out since they had the wrong gravity direction. The interaction with the actors was animated into CGI versions of the branches.

## On Location

During the shoot a laptop was used to grab video feedback and edit shots together. This allowed Jonathan to determine if he had gotten the performance he wanted in a particular shot.

## The Design Aspect

Jonathan Glazer had a very particular wall explosion in mind. The effects team had to develop the look of the explosion. The explosion was broken down into individual elements to achieve the desired final result; broken plaster pieces with wall paper, broken interior pieces, debris, several different plaster dust passes, actor trail dust, etc..

The forest needed a different design development process. References of trees were gathered on several trips to different



botanical gardens. This led to the development of a tree based on several different real trees, combining aspects of redwood trees with pines. The final master tree was created using 7 nurbs trunk sections, 120 nurbs branches, and over 500 Paint Effects strokes. The forest was laid out with rough placement grey scale trees. From a database of trees, hybrids made from the master tree, each tree was placed and rendered to create the forest. Shots would contain up to a hundred different trees.

Contributors: Andrew Daffy, Mark Nelmes, Jake Mengers, Antony Field, Robert Krupa, Jamie Isles, Alex Parkinson, Murray Butler, Stephane Allender, Helen MacKenzie, Howard Sly, Chris Syborn, Kate Hood, Rachael Penfold, Dave Ludlum

# Pop.eye: A Pop-out Video Camera System for Personal Use

Isao Mihara, Takahiro Harashima, Shunichi Numazaki and Miwako Doi  
Corporate R&D Center, Toshiba Corporation  
{isao.mihara | takahiro.harashima | shunichi.numazaki | miwako.doi}@toshiba.co.jp

A new video camera system, Pop.eye, is able to capture a real-time 3D video based on color image and reflection image. The current prototype system (Figure 1) can capture 160x120 sized images at a rate of 25 frames per second. The captured image of a stuffed toy is viewed as the “Pop-out” image frame shown in Figure 2-left. The “Pop-out” shape is constructed from the captured reflection image. Furthermore, this system is able to output an enhanced “Pop-out” movie that extracts only the target object from background by using depth-key (Figure 2-right).

There are many conventional 3D input systems, but they are unsuitable for personal use. A range finder is a 3D digitizing device for professional use, and so it is too expensive and complicated for personal use. A stereo camera system using multi-CCD camera is reasonable for personal use. However, it takes a fairly long time to capture an image because a lot of calculation is required to construct a 3D model by using a stereo matching method. So, it is difficult to capture 3D images in real-time.

Pop.eye is being developed to solve these problems. Pop.eye is a system that realizes both low-cost and real-time capturing simultaneously by using our new 3D capturing technology [Numazaki et al. 1998; Yamauchi et al. 2000]. This technology uses the principle of the reflection of light. The system emits near-infrared light by LED and a custom-made image sensor detects only reflection via the object from LED (Figure 3). Reflection from ambient light is canceled and not detected. Because this technology is simple, it works very fast and robustly. Typically, the amount of reflection has a correlation with the distance from the object. So, depth image of the object is obtained by the transformation from reflection map to depth map. To get the object’s 3D model from the obtained depth image, the system performs a calculation that rejects only local characteristics of the object while retaining global characteristics, and constructs the object’s 3D polygon model by representing each image pixel as a 3D point and each object surface as a group of quadrangle patches. After these processes, color image is mapped on the 3D polygon model as texture (Figure 4). As the result, a “Pop-out” image such as that shown in Figure 2 is captured in real-time.

In the current prototype system, two different sensors are required in order to get reflection image and color image. We plan to integrate these two sensors into a single “Pop.eye sensor” by the year 2003. Our aim is to realize a Pop.eye system that incorporates the new sensor at a less than 100 dollars. Such a Pop.eye system would be suitable for widespread personal use, since the cost would be about the same at that of a conventional personal low-cost color CCD/CMOS camera.

## Acknowledgments

A part of the ‘Pop.eye’ project is funded by the New Energy and Industrial Technology Development Organization (NEDO).

## References

- NUMAZAKI, S., ET AL., 1998. A kinetic and 3D image input device. *CHI '98 Extended Abstracts*, 237-238.
- YAMAUCHI, Y., ET AL., 2000. Gesture-Based Ping-Pong Game Using Real-Time Depth-Image Input Device. In *Proceedings of ACM SIGGRAPH 2000*, 207.

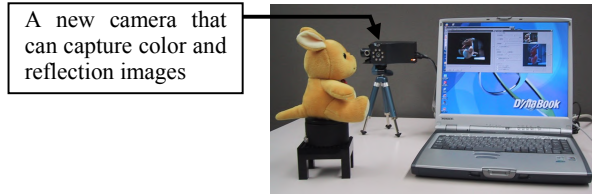


Figure 1: Overview of prototype system



Figure 2: Captured “Pop-out” video image of stuffed toy

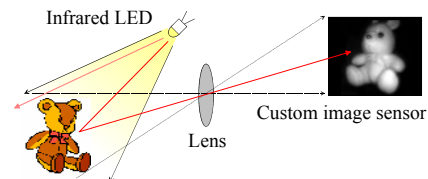


Figure 3: Principle on which capture of reflection image is based

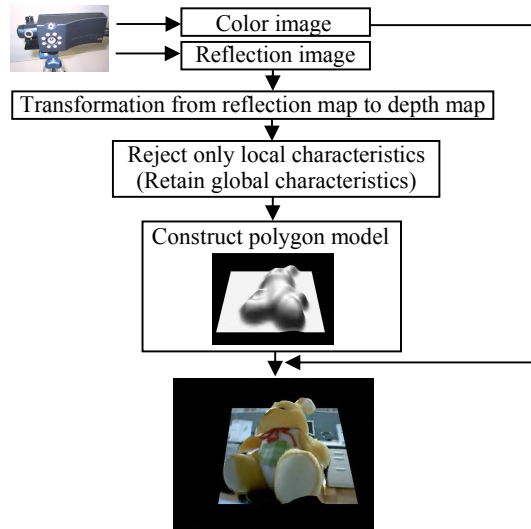


Figure 4: System architecture

# Probabilistically Placing Primitives

Adrian Secord  
Department of Computer Science  
University of British Columbia

Wolfgang Heidrich\*  
Department of Computer Science  
University of British Columbia

## 1 Stroke Placement in Image Space

Non-photorealistic rendering often requires placing drawing primitives onto a 2D canvas in such a way that the resulting tone approximates that of a greyscale reference image. Several iterative methods have been used where each stroke is tentatively placed on the canvas and the resulting tone is evaluated with respect to the reference image. [Salisbury et al. 1994; Salisbury et al. 1997; Praun et al. 2001] If the stroke over-darkens the output image it is rejected, otherwise it is accepted. While this back-and-forth iteration between the output and the reference image is capable of producing high-quality results, it is extremely costly in terms of computation and memory references.

However, if we view the reference image as a 2D probability density function (PDF), we can generate a set of primitive locations according to the PDF which will preserve tone *a priori*. Rendering then consists of taking a sequence of precomputed, uniformly distributed points in 2D and redistributing them according to the PDF. The derivation of the PDF from an image can be done rapidly to enable interactive frame rates. [Secord et al. 2002]

## 2 Generating Random Points With a PDF

Given a 2D probability density function  $p(x, y)$ , and a set of uniformly distributed random points  $\{\mathbf{p}_i\}$ , the process of redistributing the points according to  $p(x, y)$  is well-described in the literature (e.g. [Pitman 1992]).

In stippling, hatching and other applications, the samples that we would like to place have finite area. It is therefore possible that two randomly placed samples will overlap. Thus the area covered by all samples does not change linearly with the number of samples. It is possible to correct for this fact by adapting the underlying distribution function. We will consider the simplest case where all primitives cover a single pixel, and the more general case follows directly.

Let  $I(x_i)$  be the intensity of the base image at pixel  $x_i$  as a value between 0 and 1. The stippling process should create a binary image where the probability that a pixel  $x_i$  is a function of the intensity of that pixel in the base image. After placing  $N$  samples according to the density function  $p(x)$ , the probability that pixel  $x_i$  has *not* been set is  $(1 - p(x_i))^N$ . Therefore, we require that  $1 - (1 - p(x_i))^N = I(x_i)$ , or  $p(x_i) = 1 - \sqrt[N]{1 - I(x_i)}$ .

Thus,  $p(x_i)$  can be computed if we know the number of samples  $N$  that need to be placed. This number is determined uniquely by enforcing that  $p(x_i)$  is a probability density function, that is, by demanding that  $p(x_i)$  sums to one over all pixels.

For general images  $I(x_i)$ , this equation system cannot be solved directly. Fortunately it is possible to find the right  $N$  via a binary search between a maximum and a minimum value.

## 3 Results

The included images show the results of our algorithm when tied to a general primitive-drawing engine. Both images can be produced



Figure 1: A 3D model of foot bones, rendered to an image and drawn with small stipple primitives.

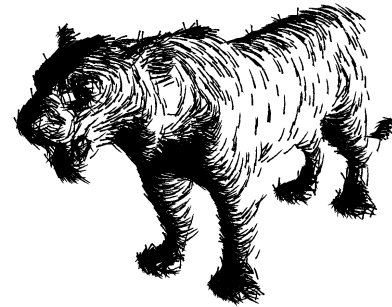


Figure 2: A 3D model of a tiger, rendered to an image and drawn with stroke primitives.

in real-time at 15 frames per second, including the rendering of the primitives.

## References

- PITMAN, J. 1992. *Probability*. Springer.
- PRAUN, E., HOPPE, H., WEBB, M., AND FINKELSTEIN, A. 2001. Real-time hatching. In *Proceedings of SIGGRAPH 2001*, 579–584.
- SALISBURY, M., ANDERSON, S., BARZEL, R., AND SALESIN, D. 1994. Interactive pen-and-ink illustration. In *Proceedings of SIGGRAPH 94*, 101–108.
- SALISBURY, M., WONG, M., HUGHES, J., AND SALESIN, D. 1997. Orientable textures for image-based pen-and-ink illustration. In *Proceedings of SIGGRAPH 97*, 401–406.
- SECORD, A., WOLFGANG, H., AND STREIT, L. 2002. Fast primitive distribution for illustration. In *Proc. 13th Eurographics Rendering Workshop, Pisa, Italy, June 26 – June 28, 2002*. To appear.

\*{ajsecord,heidrich}@cs.ubc.ca



# Probability Paint: Controlling Group Characteristics with PDFs

Kathleen Gretchen Greene  
ggreene@alumni.princeton.edu

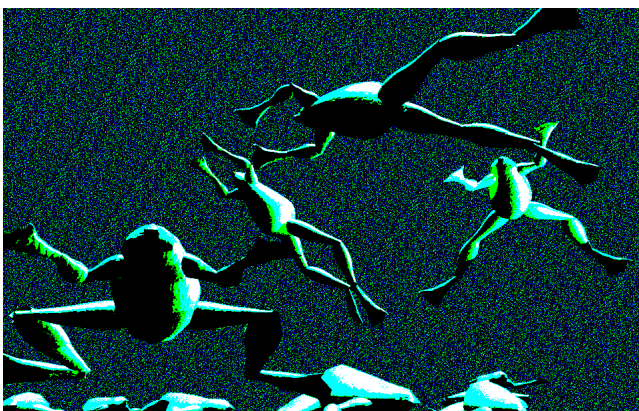
A plague of frogs, a river of stones, a dragon's back of scales... When a large group should exhibit variation in some characteristic, how can we describe and experiment with that variation and the look and feel of the group as a whole?

## 1. Shapeable Probability Distribution Functions

Shapeable probability distribution functions (SPDFs) together with a "probability paint" interface provide a powerful, intuitive tool with broad applications in high-end production. Pixar used particles and user shapeable PDFs to create stylized rain in "A Bug's Life". PDI/Dreamworks used probability and fashion rules to outfit characters in their "Shrek" crowd scenes.

These ideas can be extended to create a framework that allows high level control of any characteristic of a group, while retaining the capacity to directly control individuals. For example, we might control the heights of jumping frogs or the feel of frenetic energy, the amount of squash and stretch or the pause length on the ground. We might control color or size or pattern - anything that can be specified numerically, where numerically need not mean completely procedural. It might mean an enumeration of hand keyed animations or painted textures.

SPDFs are most useful if there is a continuum of acceptable values or an intuitively well ordered set of say, at least ten discrete values. But, the framework still works if the choices are few in number and/or not intuitively well ordered. An SPDF is a visually intuitive curve describing what proportion of a population has each given value for the given characteristic. In the very simplest case it could mean something as simple as "20% red, 80% black." But as the number of possibilities increases, it quickly becomes cumbersome and even impossible to individually list and edit the percentages.

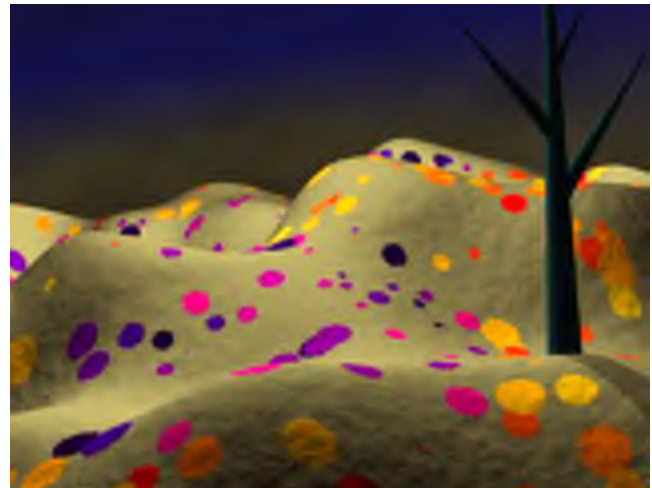


Jumping Frogs. SPDFs control pause length, jump height and jump direction in a group of 50 frogs.

## 2. Probability Paint

Once the user begins to find applications for SPDFs, similar applications can be grouped together into "probability palettes" and the different SPDFs can be applied and blended in different regions with "probability paint".

For example, one palette might consist of SPDFs representing color. A quick color study can suggest broadly what the general color impressions of a field of flowers should be. We then choose a few SPDFs from our palette to assign to the colors used in the color study and make a 2D map (or possibly paint directly on 3D geometry) to assign the SPDFs and final colors.



Field of flowers. Probability paint and SPDFs control spot size and color. Use as is for a simple cartoony landscape or use spot color to drive a procedural flower generator.

It's easiest to visualize the use of probability paint in procedural shading, to delineate areas of color or pattern, but it can be used as a high level interface anywhere an SPDF can be used, to add controllable variation in character animation, modeling, shading and effects.

### References

Herman, Daniel L. 2001. RainMan: Fluid Pseudodynamics with Probabilistic Control for Stylized Raindrops. *Conference Abstracts and Applications ACM SIGGRAPH 2001*, ACM Press/ACM SIGGRAPH, Computer Graphics Proceedings, Annual Conference Series, ACM, 246.

Modesto and Rangaraju 2001. Generic Character Variations, "Shrek" The Story Behind the Screen. Conference Lecture, ACM SIGGRAPH 2001, Los Angeles.

# Pushing the limits of L-systems for time-lapse vine growth in “The Time Machine”

Jonah Hall  
Digital Domain



Figure 1. A final composite of the L-system vines and flowers

## Background

Dreamworks Pictures film “The Time Machine” presented a number of interesting problems related to creating time-lapse photography using computer animation. I was assigned with the task of making realistic vines creep and grow along the surface of the green house as the main character began to move forward in time. In order to keep such a complicated task manageable, I needed to organize the order in which things were processed to keep things workable as changes were made. The pipeline needed to be designed in such a way so that the maximum amount of processing is taken care of at the earliest stage.

The vine growth animation for Time Machine required me to generate a unique and manageable growth system and a rendering system that complimented this movement. The work was completed with Side Effect Software’s Houdini and Pixar’s Photorealistic Renderman.

## Motion

Vine motion was broken into two categories designated pre-cook and post-cook. The procedural animation of generic vine models was separated into the pre-cook phase. These were basically individual vines grown straight up from world center which had all of their individual characteristics baked in. The term “pre-cook” refers to the fact that, because the generation of this animation was so computationally expensive, the animation itself was done only once and then saved and applied multiple times within the scene later.

Thus, post-cook motion is any animation done to the vines after reading those sequences of cooked vines off disk. This motion happened on a per frame basis so that individual frames could be rendered on the render farm without any simulation roll-up costs.

## Pre-cook motion

The Houdini L-systems operator was excellent for sculpting the look of a plant, but was limited in its ability to adjust values to cause motion over time. An alternative approach to create animated motion on the L-system would be to do all motion of the vines in a procedural network after the L-system has cooked. This was impossible since, after exiting the L-system environment, you leave behind all hierarchical information. A branch no longer knows where it grew from, so rotating the data along the bases to move them like kinematics becomes impossible.

The first thing we did was to feed each vine a different generation value so that they had unique growth behavior. A simple procedural curve network was used to create a number of different

growth patterns. Each vine builds itself on one of these curves, selected by the index of the vine. Thus, variable curve values allow for the appearance that branches are overtaking one another by increasing and decreasing speed at different times.

One of the challenges was to insert some behavior into the rule that gave the look of constant breathing of the branches. I wanted to make the branches look as if they had constant jitter going on to create a time lapse effect. To achieve this, I placed little rotation “motors” all over my vine joints. Every generation found its rotation rule and randomly rotate/pivot/twist the orientation of each branch. This caused my branches to jitter a bit every frame, and often slide around a little over time if probability leaned in either the positive or negative direction over time. Since this was a recursive rule, as long as the vine gained new generations, the rotation values would always be changing, causing twig fluttering.

## Post-cook motion

The vine animation along the side of the greenhouse needed to be carefully timed with the camera pan. Since there was a set number of pre-cooked frames of each vine’s growth, I was able to start each vine growing exactly when I wanted it to. The Director asked to have individual vines be shorter or taller at different times, so I needed a system to be able to offset the start frame of each vine as needed. To achieve this, I imbedded an offset into the height of each birth point of the template geometry.

Using the ability to assign varying parameters to different copies of the same geometry, I was able to read in the sequence of vines along a series of points. By modeling the height of the template points slightly, I was able to interactively time the vine growth so that they didn’t grow out of camera, didn’t start or finish too early, or any other timing related problems.

## Vector Blur

Motion blur was used to create look that indicated that the vine had repositioned itself a number of times while the shutter was open for the duration of a frame. Houdini allowed me to fabricate vectors indicating motion blur that didn’t necessarily correspond to the motion taking place in the animation. The vectors were created by adding together two directions to each point per frame. The first value represented the direction that the point was actually traveling in during the growth.

The second value was generated by rolling 4d noise vectors based on position through the points of the vine, giving me smooth, flowing blur around the vine. Adding these values gave me the direction of the motion blur for each portion of the vines.

Given those directions, I still needed to make sure that the blur did not “push” the geometry around, but only blurred it. I backed the position of every point halfway along the vector that that frame would be blurring, assuring that the geometry would always be centered along the blur and not appear to pop every frame.

This magic combination of steps allows me to create smeared, “time lapse” looking blurs across different parts of the vines at different times. This technique was used in varying degrees on most of the plants in Time Machine.

# Quikwriting on the Responsive Workbench

Jerome Grosjean, Sabine Coquillart\*  
i3D INRIA FRANCE

## 1 Text Input in Virtual Environment

Virtual environments (VEs) like the projection-based Responsive Workbench have greatly enhanced the interactive visualization and manipulation of 3D objects. In these configurations, the classical desktop interaction techniques have to be reconsidered. In particular, a simple operation like text typing, although needed for basic operations like saving one's work under a specific file name or entering a precise numerical value inside an application, is still problematic.

The first naive solution to let the user enter text through the classical keyboard is not practical. A VE user is often equipped with more or less cumbersome interaction devices, like data gloves or a tracked stylus in the hand. Forcing a user to exit the 3D world and get the physical keyboard to type a short text/data/command is not desirable. Fastening a miniature keyboard on a user's arm is not always convenient too: it is a good solution when typing text is a predominant task and when the user's hand is free for typing. At the opposite wearing a keyboard on the arm can be annoying and typing can be complicated if the user is wearing other devices.

Indeed the search for the best interaction paradigm for text typing in VE depends strongly on the current devices held in the hands or worn. In an environment where users are working with data gloves, the Pinch Keyboard of Bowman[Bowman et al. 2001] or VType developed by Evans et al.[Evans et al. n. d.] constitutes interesting approaches. On the Responsive Workbench configuration, the stylus, a spatially tracked pen device with a button, is commonly used in the dominant hand to interact with the virtual objects. A better solution here is to develop an interaction technique based on this device. A virtual keyboard could be used, with the stylus clicking on the keys. However the typing would remain very slow like when someone types with only one finger and would require precision to select the keys.

We notice that the well-known personal digital assistants (PDA) have a similar requirement : a pen-based interaction. One specific writing method on PDAs caught our attention and we decided to experiment it on the workbench.

## 2 Quikwrite

Quikwrite[Perlin 1998] is a quick writing paradigm for the Palm Pilot<sup>TM</sup>. It consists of a square area divided in 3 by 3 little square slots and a stylus. It lets the user write without ever picking the stylus up off the surface. Each letter writing starts from the central slot and is done by moving the stylus to one or more adjacent slots and then coming back to the center. The combination of the slot reached when leaving the central slot and the one the stylus comes from when finishing the movement determines which letter is written. Writing is thus done through a sequence of loops crossing the central slot. In our porting of quikwrite on a two-screen workbench (Fig. 1) the square surface is floating in the 3D scene. A tracked stylus is used for the interaction and is activated by pushing once the button. It gives the opportunity to stop writing at anytime by pressing the button again. Translations of the stylus are mapped to a virtual pen sticked to the surface. Restricting movements to two degrees of freedom decreases the required precision when writing.

The virtual pen gives a visual feedback to the user who is not directly writing on the surface. The text is displayed above the surface as a 3D object (Fig. 1).

## 3 Feedback

We conduct an informal evaluation with 5 novice users. All were familiar with both computers and the workbench. They were first explained how quikwrite works. Then we let them do some basic manipulations and write some sentences to train them with this new writing tool. At first they wrote slowly, due to their ignorance of the key positions, but they quickly tend to accelerate on some simple string of keys. One of the users found that having a vertical orientation for quikwrite's surface was tiring. A better orientation could be to put quikwrite horizontally or slightly steep. Another user noticed that when he moved into a wrong slot there was no method to cancel his action, except by ending his movement in the central slot and choosing backspace as the next character. He suggested to let the pen leave the surface in this case. All users were pleased by this writing method at the end of their training period. We plan in the future to do a more complete and precise evaluation of the quikwrite tool, especially in comparison with other stylus-based writing methods like virtual keyboards. The addition of a transparent glass plate or palette is also considered to give an haptic feedback and a support for the wrist.

**Acknowledgements :** This work has been partially funded by the RNTL Inventor Immersif project. The workbench used has been partially funded by the Regional Council of Ile de France.



Figure 1: Quikwrite on the Responsive Workbench

## References

- BOWMAN, D., LY, V., AND CAMPBELL, C. 2001. Pinch keyboard: Natural text input for immersive virtual environments. Tech. Rep. TR-01-15, Virginia Tech Dept. of Computer Science.
- EVANS, F., SKIENA, S., AND VARSHNEY, A. Vtype: Entering text in a virtual world. submitted to International Journal of Human-Computer Studies.
- PERLIN, K. 1998. Quikwriting: Continuous stylus-based text entry. In *Proceedings of the ACM Symposium on User Interface Software and Technology*, Fast Pen Input, 215–216.

\*e-mail: Jerome.Grosjean@inria.fr, Sabine.Coquillart@inria.fr



# RAPTOR: Towards Augmented Paleontology

Oliver Bimber and L. Miguel Encarnação  
Fraunhofer Center for Research In Computer Graphics  
{obimber,me}@crcg.edu

Paleontology is filled with mysteries about organisms such as plants and animals that lived thousands, millions, and billions of years before the first modern humans walked the earth. To solve these mysteries, paleontologists rely on the excavation, analysis, and interpretation of fossils. Fossils are the remains or traces of ancient life forms that are usually preserved in stones and rocks. Examples include bones, teeth, shells, leaf imprints, nests, and footprints. Such fossil discoveries reveal what life on our planet was like long ago. Fossils also disclose the evolution of organisms over time and how they are related to one another. While fossils reveal what ancient living things looked like, they keep us guessing about their color, sounds, and most of all their behavior. Each year, paleontologists continue to piece together the stories of the past.

The RAPTOR project aims at supporting both –paleontologists and us, the visitors of museums around the world– to explore and understand these stories in an exciting and effective way. By bringing together the oldest artifacts that once have been put to sleep by nature, and the latest technology that has been developed by science, our goal is to wake them up after millions of years.

The key technology for our approach is a new multi-user Augmented Reality display – the Virtual Showcase [Bimber et al. 2001]. The Virtual Showcase is a projection-based and high-resolution display that offers an innovative way of accessing, presenting, and interacting with scientific and cultural content. Conceptually, the Virtual Showcase is compatible with conventional showcases used, for instance, by museums. Technologically, however, it allows the display of computer generated stereoscopic 3D graphics and animations together with real artifacts within the same space.

The presented images illustrate the latest results of our project. A physical skull of a mid-cretaceous raptor –a *Deinonychus* (‘terrible claw’)– is displayed within the Virtual Showcase. A high-precision data set of the skull’s geometry has been captured using state-of-the-art 3D scanning technology. This model is used to create view-dependent illumination effects directly on the skull’s surface. Missing parts, such as muscles, additional bones, and the skin have been reconstructed using off-the-shelf 3D modeling tools. Figure 1 illustrates all scanned and modeled data sets. Figures 2 and 3 reveal the view on the augmented physical skull as seen by one of the observers. In Figure 2, the reconstructed muscles and the bones are visualized after registering them precisely to the bone structure. Figure 3 shows how the skull can be augmented with the virtual skin. Note, that the images are normally rendered stereoscopically to generate correct and matching depth effects.

## References

BIMBER, O., FRÖHLICH, B., SCHMALSTIEG, D., AND ENCARNAÇÃO, L.M. 2001. The Virtual Showcase. *IEEE Computer Graphics & Applications*, vol. 21, no.6, pp. 48-55, *ACM SIGGRAPH Sketches and Applications*, pp. 227.

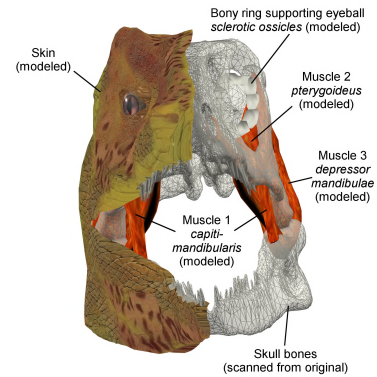


Figure 1: Different scanned and modeled data sets.



Figure 2: Physical skull augmented with virtual muscles and bones – as seen by one observer within the Virtual Showcase.



Figure 3: Physical skull augmented with virtual skin. Most of the skull is covered by the skin. The teeth, however, are real.

The Virtual Showcase project is supported by the European Union (IST-2001-28610)



# Real Time Video Effects on a PlayStation2

Sarah Witt, Sony B&P Research Labs, Jays Close, Viables, Basingstoke, Hants, RG22 4SB, UK  
sarah.witt@adv.sonybpe.com

The Sony PlayStation2, with its powerful rendering and vector processing capabilities, built-in MPEG decoder, and large capacity hard disc drive possesses fundamental assets that make it an ideal platform for processing video. Research done at Sony BPRL, where we have extensive experience of video effects processing, has shown that this inexpensive games console can, in fact, create in real time the type of video effects which normally require investment in dedicated hardware (or endless patience as a PC renders them slowly).

## 1. SYSTEM

By using the internal MPEG-2 decoder inside the Emotion Engine processor of the PS2, two concurrent video sources can be streamed from the hard disc drive and decoded. If constant bit rate MPEG-2 I-frame only video is used, the decoding can be done very quickly and in parallel to other processing. The high bandwidths available internal to the PS2 enable quick and easy transfer of full-screen video to and around the Emotion Engine and Graphics Synthesiser.

## 2. EFFECTS

The types of effects implemented range from simple fade or wipe transitions to complex nonlinear 3D transformations and combined effects, such as recreating the look of an old sepia film.

The nonlinear 3D transformations, such as pageturn, sphere and ripple, are implemented on a tiled basis, with the vertices recalculated every field, depending on the input parameters. By using look-up tables the calculations required for the nonlinear part of the transformation can be kept to a minimum. Then the linear part of the transformation can be processed very efficiently in the parallel floating point vector processors within the PS2. At this point, lighting based on the orientation of each tile can also be added, to enhance the 3D effect. The foreground video source is used as a texture and mapped onto the calculated tile positions. Anti-aliasing of the video can be implemented by way of mip-mapping. A second video source can be used as the background.

The “old film” effect uses sepia (created using the luminance data and a colour look-up table), random scratches, textured blotches and grain, as well as shake, defocus and the occasional dropped frame, to apparently “age” the video. This can be combined with wipes or fades as transitions to another video source. This is but one example of how effects on the dynamic input video streams can be combined with pre-programmed graphics and textures to produce more complex looks and enhancements to original footage.

## 3. ACKNOWLEDGEMENTS

We would like to acknowledge and thank Sony Computer Entertainment Europe developer support in London, and Sony BSNC in Atsugi, Japan.



Fig.1 Pageroll Effect

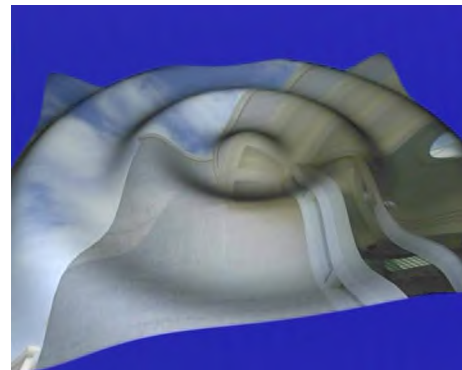


Fig.2 Ripple Effect



Fig.3 Old Film Effect

# Real-Time 3D Interaction for Augmented and Virtual Reality

Simon Prince<sup>1</sup> Adrian David Cheok<sup>1</sup> Farzam Farbiz<sup>1</sup> Todd Williamson<sup>2</sup> Nik Johnson<sup>2</sup> Mark Billinghurst<sup>3</sup> Hirokazu Kato<sup>4</sup>  
<sup>1</sup>National University of Singapore {elesp,eleadc,eleff}@nus.edu.sg <sup>2</sup>Zaxel Systems {toddw,nik}@zaxel.com <sup>3</sup>University of Washington grof@hitl.washington.edu <sup>4</sup>Hiroshima State University kato@sys.im.hiroshima-cu.ac.jp

## 1. SYSTEM OVERVIEW

We demonstrate a real-time 3-D augmented reality video-conferencing system. The observer sees the real world from his viewpoint, but modified so that the image of a remote collaborator is rendered into the scene. For each frame, we estimate the transformation between the camera and a fiducial marker using techniques developed in Kato and Billinghurst [1999]. We use a shape-from-silhouette algorithm to generate the appropriate view of the collaborator in real time. This is based on simultaneous measurements from fifteen calibrated cameras that surround the collaborator. The novel view is then superimposed upon the real world image and appropriate directional audio is added. The result gives the strong impression that the virtual collaborator is a real part of the scene.

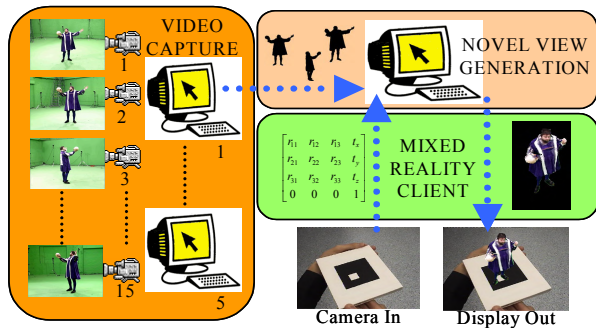


Figure 1. System overview: The 3-D transformation of the marker is determined in each incoming frame. The novel view generation system takes input from 15 cameras and generates the corresponding view of the actor, which is rendered back into the scene in real-time.

## 2. NOVEL VIEW CREATION

The crux of the system is a real-time novel view generation algorithm, based on shape-from-silhouette information (see also [Matusik et al. 2000]). The algorithm proceeds on a per-pixel basis, calculating the depth for each pixel in the desired image, and then estimating the color using a weighted average of the relevant pixels of nearby images. This approach is much faster than methods that calculate the entire shape and our system runs at 30 Hz with 384x288 resolution.

## 3. APPLICATIONS

We develop a 3-D augmented reality video-conferencing system. The system captures a large (3x3m) area, allowing full expression of gestures. The user can see the collaborator from different viewpoints by moving his head or manipulating the card marker. This system solves many problems that hinder traditional video-conferencing solutions, including the lack of any spatial cues to participants and the poverty of gestural information.

We also present the first demonstration of interaction in virtual environments with a “live” fully 3-D collaborator using the example of a guide to an art gallery. The user can view the art gallery and converse with a collaborator who is also in the virtual space. Interaction is much more natural than with a graphical avatar. It is easy and natural to refer to aspects of the environment with normal gestures and body orientation.

Finally, we consider interaction between users in the real

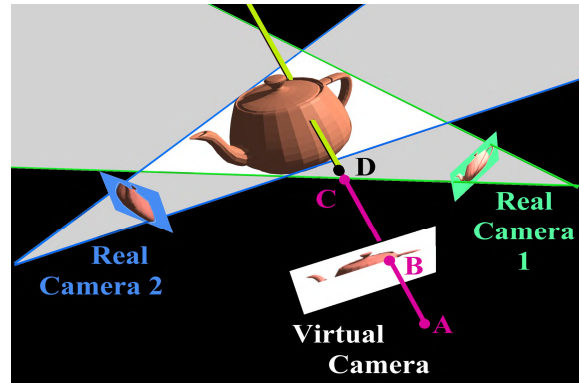


Figure 2. Novel view generation: The shape-from-silhouette algorithm divides each image into foreground (silhouette) and background. The recovered shape is the intersection of the 3-D volumes defined by each silhouette. The virtual image is constructed by projecting a ray through each pixel and finding the intersection with the object. Pixel color is determined by sampling the image of this 3-D point in nearby real cameras.

world and collaborators in a virtual space, using a “tangible” AR interface. Users can move a collaborator from one virtual environment to another using a tracked paddle. A game is developed in which the aim is to drop objects onto the collaborators head while he tries to evade them, in real time. For more details please visit <http://mixedreality.nus.edu.sg>.

### References:

KATO, H. AND BILLINGHURST, M. 1999. Marker tracking and HMD calibration for a video based augmented reality conferencing system. *In Proc. IWAR*, 85-94.  
 MATUSIK, W., BUEHLER, R., RASKAR, R., GORTLER, S.J. and MCMILLAN, L. 2000. Image-Based Visual Hulls. *In Proc. SIGGRAPH*, 369-374.



Figure 3. Applications include augmented reality video-conferencing (top-left), live avatars in virtual reality (top-right) and tangible interaction between users in augmented and virtual spaces (bottom)

# Real-Time Image-Space Outlining for Non-Photorealistic Rendering

Jason L. Mitchell  
ATI Research  
JasonM@ati.com

Chris Brennan  
ATI Research  
CBrennan@ati.com

Drew Card  
ATI Research  
DCard@ati.com

## Introduction

In Non-Photorealistic Rendering (NPR), outlines at object silhouettes, shadow edges and texture boundaries are important visual cues which have previously been difficult to generate in real-time. We present an image-space technique which uses pixel shading hardware to generate these three classes of outlines in real time. In all three cases, we render alternate representations of the desired scene into texture maps which are subsequently processed by pixel shaders to find discontinuities corresponding to outlines in the scene. The outlines are then composited with the shaded scene.

## Outlining Silhouettes and Creases

For the cases of silhouette and crease outlining, we render the full scene's world-space normals and eye-space depths into an RGBA ( $nx_{world}, ny_{world}, nz_{world}, depth_{eye}$ ) texture using a vertex shader to populate the color channels [Decaudin1996; Saito and Takahashi 1990]. This texture is subsequently processed using a discontinuity filter implemented as a pixel shader and composited over the shaded image. With this technique, creases and silhouettes both produce discontinuities in the image and there is no real distinction between them. Note that we are not processing the contents of the depth buffer, but just happen to be generating a texture channel that has a semantic meaning that is the same as the depth buffer (albeit at lower precision).

## Outlining Shadows

To outline shadow borders, we use the same renderable texture and image filtering technique as were used for silhouettes and creases above. In the case of shadows, we use geometric shadow volumes and the stencil buffer to flag pixels which are in shadow for a single light source. For these pixels, we negate the world-space normal before it is written to the renderable texture. This creates additional discontinuities in the  $\{nx_{world}, ny_{world}, nz_{world}, depth_{eye}\}$  image which align with the shadow boundaries as shown in Figure 1a. Multiple lights can be handled with multiple  $\{nx_{world}, ny_{world}, nz_{world}, depth_{eye}\}$  images and composited with the shaded image.

## Outlining Texture Boundaries

Texture boundaries lie between areas of an object which are textured differently but which do not generate a geometric discontinuity. To handle this case, we use a technique similar to priority shadow buffers [Hourcade 1985]. The regions of different texture are colored with IDs which are suitably different from adjoining regions (Figure 1c). This creates a discontinuity in the ID image. We have found that applying an ID map of insufficient resolution can lead to edge detection issues due to filtering of the ID texture. Specifically, extreme magnification and bilinear filtering of an ID texture can lead to poor detection of the texture boundary edges. To combat this, we have used relatively large textures with linear-mip-nearest filtering.

## Enhancing Outlines with Morphology

To further allow us to customize the look of a scene, we can choose to dilate or erode the edge image prior to compositing it with the shaded scene. This is done with dilation or erosion pixel shaders [Mitchell 2002]. Figures 1b, 1d and 1f all contain outlines which have been dilated.

## Optimization

It is possible to reduce the cost of compositing the edge images with the shaded scene using alpha testing to mask out non-edge pixels. Since there are typically a large number of non-edge pixels in a scene, as seen in Figures 1b and 1d, this can be a significant savings. This optimization can also be applied when dilating the edge images.

## Future Enhancements

In the near future, we expect to be able to render to multiple textures at the same time, going even farther toward a general g-buffer [Saito and Takahashi 1990]. This will allow us to collapse many of our rendering passes into a single rendering pass into multiple textures. Additionally, we have not addressed issues of outline antialiasing. While we could have just made our edge detection and morphology filters "fuzzy," this is probably an area that needs more focused and formal research.

## Conclusion

We have applied hardware pixel shaders to outlining problems by rendering alternate representations of our scenes into texture maps. By processing these rendered images with edge detection filters, we are able to extract outlines which can be composited with a shaded scene in real-time.

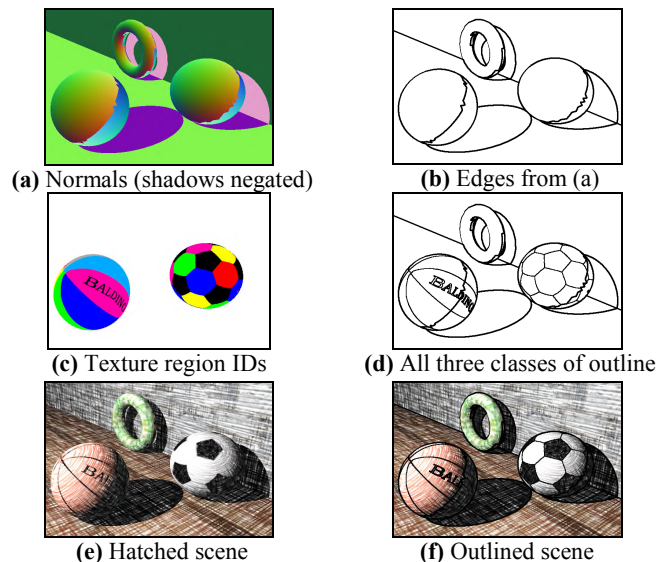


Figure 1 – Intermediate results and final outlined scene

## References

- DECAUDIN, P. 1996. Cartoon-looking rendering of 3D-scenes. Technical Report INRIA 2919, Université de Technologie de Compiègne, France.
- HOUSCADE, J. C. AND NICOLAS, A. 1985. Algorithms for Antialiased Cast Shadows. In *Computer and Graphics*, 9, 3, 259-265.
- MITCHELL, J. L. 2002. Image Processing with Direct3D Pixel Shaders. In *Vertex and Pixel Shader Tips and Tricks*, Wolfgang Engel editor, Wordware.
- SAITO, T AND TAKAHASHI, T. 1990. Comprehensible Rendering of 3-D Shapes. In *Proceedings of SIGGRAPH 1990*, ACM Press / ACM SIGGRAPH, New York. 197-206



# Real-Time View Synthesis Using Commodity Graphics Hardware

Ruigang Yang, Greg Welch, Gary Bishop, Herman Towles \*

Department of Computer Science, University of North Carolina at Chapel Hill

## 1 Our Real-Time View Synthesis Method

We present a novel use of commodity graphics hardware that effectively combines a plane-sweeping algorithm [Collins 1996] and view synthesis in a single step for real-time, on-line 3D view synthesis. Unlike typical stereo algorithms that use image-based metrics to estimate *depths*, we focus on using image-based metrics to directly estimate *images*. Using real-time imagery from a few calibrated cameras, our method can generate new images from nearby viewpoints, without any prior geometric information or requiring any user interaction, in real time and on line.

For a desired new view  $C_n$ , we discretize the 3D space into a number of candidate focal planes  $\{D_i\}$  parallel to the image plane  $C_n$  of the desired view. These candidate planes discretize each desired view ray into a finite set of sample points. We step through the set of candidate focal planes, looking for the sample point along each view ray that offers the maximum color consistency among input images. Each final image pixel is determined by the color at this point of maximum consistency.

To accomplish this, for each candidate plane  $D_i$  we project (texture) the input images onto that plane. We then render the resulting textured plane onto the image plane of  $C_n$  to get an image ( $I_i$ ) of  $D_i$ . We combine these two operations into a single homography (planar-to-planar) transformation. In the first row of Figure 1, we show a number of images from different planes. Note that each of these images contains the projections from all input images, and the area corresponding to the intersection of objects and the correct candidate focal plane remains sharp. For each pixel location ( $u,v$ ) in  $I_i$ , we compute the mean and Sum of Squared Difference (SSD) score. The final color of ( $u,v$ ) is the color with minimum SSD score in  $\{I_i\}$ .

We have discovered a novel use of graphics hardware to carry out the *entire* computation on the graphics board. Modern graphic cards, such as NVIDIA's GeForce series, provide a programable means for per-pixel fragment coloring through the use of *register combiners* [Kilgard 2000]. We exploit this programmability, together with the texture mapping functions, to carry out the entire computation on the graphics board.

## 2 Hardware Acceleration

In our hardware-accelerated renderer, we step through the candidate focal planes from near to far. At each step ( $i$ ), there are two stages of operations, *scoring* and *selection*. In the *scoring* stage, we set up the transformation according to the new view point. We then project the reference images onto the plane  $D_i$ . The textured  $D_i$  is rendered into the image plane (the frame buffer). In this stage, the *Pixel Shader* is configured to compute the SSD score and the frame buffer acts as an accumulation buffer to keep the mean color (in the RGB channel) and the SSD score (in the alpha channel) for  $D_i$ . In the second row of Figure 1, we show the SSD score images (the alpha channel of the frame buffer) at different steps.

In the next *selection* stage, we need to select the mean color with the smallest SSD score. The content of the frame buffer is copied to a temporary texture ( $T_{work}$ ), while another texture ( $T_{frame}$ ) holds the mean color and minimum SSD score from the previous step. These two textures are rendered again into the frame buffer through

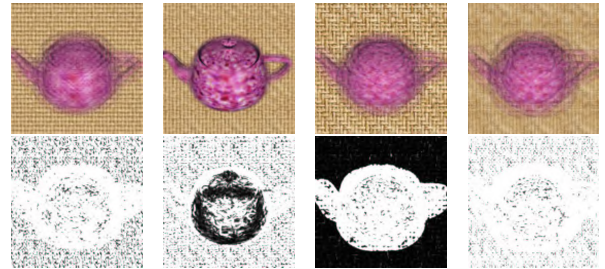


Figure 1: Mean color (above) and SSD scores encoded in the alpha channel (below) for different candidate focal planes in the first *scoring* stage. The scene consists of a teapot and a textured planar back wall.

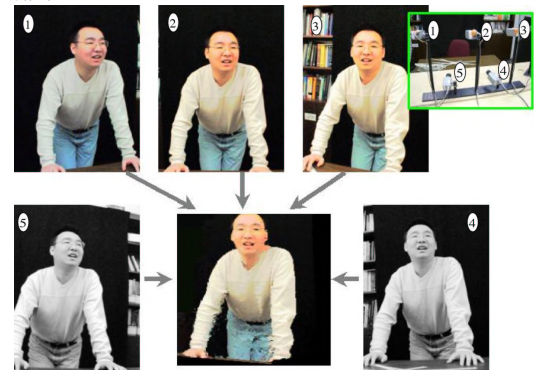


Figure 2: Example setup for 15 frame-per-second on-line reconstruction using five cameras.

orthogonal projection. We reconfigure the *Pixel Shader* to compare the alpha values on a per pixel basis, the output color is selected from the one with the minimum alpha (SSD) value. Finally the updated frame buffer's content is copied to  $T_{frame}$  for use in the next step.

One major advantage of our method is that once the input images are transferred to the texture memory, all the computations are performed on the graphics board. There is no expensive copy between the host memory and the graphics board, and the host CPU is essentially idle except for executing a few OpenGL commands.

## 3 Preliminary Results

We have implemented and tested our method with an NVIDIA GeForce3 graphics card. A typical reconstruction (256x256 output resolution with 50 candidate planes) takes less than 70 ms. Figure 2 shows our setup and a sample result. More compelling, interactive results can be found in the accompanying videotape. To better illustrate the results we include some live red-blue stereo sequences on the video.

## References

- COLLINS, R. T. 1996. A Space-Sweep Approach to True Multi-image Matching. In *Proceedings of Conference on Computer Vision and Pattern Recognition*, 358–363.
- KILGARD, M. J. 2000. A Practical and Robust Bump-mapping Technique for Today's GPUs. In *Game Developers Conference 2000*.

\*E-mail: {ryang, welch, gb, towles}@cs.unc.edu



# Recent Exact Aesthetics Applications

Tomáš Staudek, Petr Machala\*

Faculty of Informatics, Masaryk University of Brno, Czechia

E-mail: {toms | xmachala}@fi.muni.cz

Exact aesthetics is a challenging field of the computer-aided visual creativity, reconstructing the methods of design and criticism on an algorithmic basis and integrating a computer into processes of an artistic creation and aesthetic evaluation. The discipline involves principles of mathematics, geometry, theory of communication, perceptual psychology, computer graphics, or generative arts into classifying and assessing the aesthetic phenomena. The sketch introduces recent applications in this domain.

The *Arthur* program [Machala and Staudek 2002] is a rule-based aesthetic system for designing abstract images, textures and patterns — a subset of which can be appraised as an “algorithmic art”. The system is inspired by formal theories of aesthetic reasoning [Stiny and Gips 1978], utilizing rewriting grammars to produce constructivist images. Besides the grammar, it is a repertory of symbols that particularizes the final design. The aesthetic system is remarkably open to users’ creative aptitude; the user — an artist or a designer — can import ready-made construction elements and define an appropriate grammar, or let the program create its own symbolism. The symbols are placed over the canvas algorithmically by performing executive rules of the respective grammar; the design can also be “sketched” manually in an editor and made up afterwards with an algorithm. The program controls a course of the image generation by a wide range of evocation criteria concerning the symbols’ arrangement, coloring, contrast or appearance preferences.

It is perhaps the most noteworthy component of the program which enables an instant aesthetic evaluation of images. The “beauty” of the resultant pattern is determined from its structure (symbol arrangement) and information content (symbol selection) according to the measures of unity, complexity, harmony, variety, entropy or redundancy [Klinger and Salingeros 1999]. The program enumerates whether the design is disturbing or monotonous, what amount of information is assigned to individual symbols, what time it takes to comprehend the entire image, how much attractive the pattern can be considered, *et cetera*. The process of an algorithmic creation and evaluation may be then iterated until a certain level of the aesthetic metrics’ combination is reached. Figures 1 and 2 show examples of different images made of identical symbols together with their evaluation.

A cooperation with the Military Academy of Brno, CZ, has been established recently to assist designers of the “augmented reality” equipment for fighter pilots. The research employs the evaluative unit of the introduced aesthetic system in examining the headpiece displays’ readability. Classifying the flight data with respect to their importance, suggesting an appropriate iconicity and proposing visual structures that will exploit yet not exhaust the pilots’ communication channels are the key tasks in this project. Since the research has only started, it is too soon to anticipate its development and results presently.

Considering future work, *Arthur* will be enriched with additional aesthetic functions for evaluating proportions, cohesion, homogeneity, or rhythm of patterns; there are over dozen of structure

metrics to be implemented. A collaboration with psychologists is now in progress since the proposed measures require better psychometric interpretation.

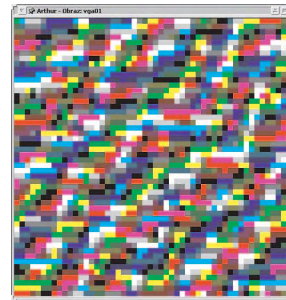


Figure 1: A simple image carrying 11 Kb of information with 2% of redundancy. The pattern is evaluated as complex, low-ordered yet considerably attractive.

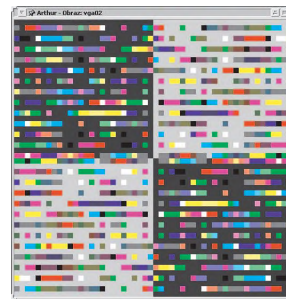


Figure 2: A similar image carrying 8 Kb of information with 32% of redundant symbols. The pattern is appraised as structured, well-ordered and monotonous.

## References

- KLINGER, A., AND SALINGAROS, N. A., 1999. A pattern measure. Centre for Advanced Spatial Analysis, University College London, Pion, Ltd. URL: <http://www.math.utsa.edu/sphere/salinger/-PatternMeasure.html>.
- MACHALA, P., AND STAUDEK, T., 2002. *Arthur*; an algorithmic system for a computer-aided generation of abstract art. Project’s URL: <http://fosforos.fi.muni.cz/arthur/>.
- STINY, G., AND GIPS, J. 1978. *Algorithmic Aesthetics: Computer Models for Criticism and Design in the Arts*. University of California Press, Berkeley, Cal.

\*Acknowledgements: The work was supported by Grant Agency of Czech Republic, Contract No. GACR 201/98/K041, and by Czech Ministry of Education, Contract No. VZ MSM143300003.

# Reconstructing or Inventing the Past, A computer simulation of the *unbuilt* church by Alvar Aalto

by Andrzej Zarzycki [zarzycki@alum.mit.edu](mailto:zarzycki@alum.mit.edu)



There is a certain mystery surrounding the *unbuilt* projects or unrealized ideas of famous architects. Often there is an expectation of deeper meaning and hidden genius present in unfulfilled buildings. Some critics go as far as to claim that the best and most interesting projects remain unrealized because of the progressiveness of the ideas associated with those buildings.

Whether we agree or not with this point of view, most would admit that there is great intellectual value to be derived from studying and (re)creating *unbuilt* buildings. Designers operate within an intellectual continuum that is evident in

their work. This continuum is successively redefined and the designer's vision transcends with each consecutive project. In this sense, *unbuilt* projects are 'missing links' and serve as 'stepping stones' in an architect's creative development. Studying *unbuilt* architecture allows one the ability to trace the origin of concepts and persistence of thoughts.

The primary reason for creating virtual models of *unbuilt* buildings with the tools of computer graphics is that drawings and physical models for these building do not reveal the full meaning and potential of the crafting of form, light, and materials into powerful, resonant space. The virtual model is the only medium, which can deliver an experience compatible with real life observation.

## Principle of uncertainty

Whenever we try to restore ancient ruins, to interpret a piece of art or to translate a poem, we come across a dilemma: to what extent do we (re)create the original artifact through our own preconceived ideas. To what extent does the Present inform the Past? I was aware of this dilemma during my work and set up a methodology for addressing it. During this workshop, I would like to present my research, methodology and (re)creative process for the unbuilt Altstetten Parish Church in Zurich designed by Alvar Aalto. I would also like to share my regarding what I feel can be learned from making virtual buildings.

## Process & Methodology

Specifically, I would like to talk about the decision process used during modeling and rendering *unbuilt* buildings from the architect's original drawings. The original drawings were essential in portraying interior spaces and tectonics. However, they did not address the details of the project nor did the hand drawings describe finishes: colors, materials and textures. These limitations were overcome through research and the extrapolation of details, materials and furnishings from other buildings built by Alvar Aalto.

A critical part of the success of this project was my methodology. I divided the project into phases, each phase had its own a distinct character and objective: the *research phase* to gather reference materials (e.g. obtaining large scale prints of original drawings from the Aalto's Foundation in Helsinki, Finland) and to analyze the final state of the design; the *design phase* to resolve contradictions and missing elements through the process of extrapolation from built works and archival materials; the *modeling phase* where 3-D digital models were constructed; the *visualization phase* where ma-

terials and lighting simulations were developed, and the *presentation phase* where various simulation techniques were evaluated to best communicate the experience and the essence of *unbuilt* space.

## Not a single, but all possible designs

By going through the process of virtually constructing an *unbuilt* building one not only gains an insight into the designer's thought process but one also gains the ability to recognize all the other possible designs which could have been realized. This exercise unfolds pictures of possibilities. Since the reconstructed image is one of many possible, albeit likely the most probable interpretation of unbuilt space, the (re)creative process requires us to think about an idea about all possible spaces that can be interpreted from the same set of assumptions. What ... if, is a constant question. Changing of materials, spatial elements or site orientation would alter the space giving us a new design with different sun exposure. Testing these possible designs against the most probable design would enable us to better judge the original ideas.

Finally, the very process of creating virtual worlds and imagery provokes new and interesting questions, which may be utilized to guide designers' future efforts.

*What are the rules of resurrecting unbuilt projects and presenting them with photo realistic images?*



# Regeneration of Real Objects in the Real World

Hiroto Matsuoka\*

Akira Onozawa

Hidehiko Sato

Hisao Nojima

NTT Lifestyle and Environmental Technology Laboratories

## 1 Introduction

The vision of this work is make it possible to regenerate real objects that had existed at some moment in the past and/or at some remote location as if they have been transported to the present across space and time. The objects could be museum pieces or items in stores, for example. In this work, we have developed a quick and fully automated system that can capture a three-dimensional image of real objects. This success has brought us close to realizing our vision.

## 2 Technological Innovations

The system we have developed provides a scene in real time, in which transported real objects can exist naturally in the present real space. The system consists of a capturing device, sensor devices, and a viewer.

### 2.1 The fully automated 3D capturing system

We have developed an image-based modeling system that can automatically create CG data of high fidelity in only an hour (Fig. 1). First the system automatically captures images of an object from every direction under various lighting conditions. Then when the object is rendered, the image taken from the angle most closely approximating the direction the object is viewed from is used, and a 3D virtual space is constructed. The critical key for fully automating the process is to extract just a portion of the many photographs depicting the object of interest. Traditionally, an object is extracted by using a binary mask to determine whether each pixel in a photograph is object or background. However, if there are reflections on the surface of the objects, then it becomes difficult to automatically differentiate the object from background, and errors in extracting the object occur. We have overcome this problem by substituting a 256-tone multi-level silhouette for the conventional binary mask. With this method, those parts that cannot be distinguished as either background or object are represented by an ambiguous value. Naturally, those portions of the final CG image will also be ambiguous, but humans have an innate ability to infer the shapes of objects, so the image appears entirely natural to the human eye [Matsuoka et al. 2002]. Just by viewing the reflections that would otherwise be generated if the ambiguous parts were actually present, one can perceive images depicting glossy objects that are extremely realistic.

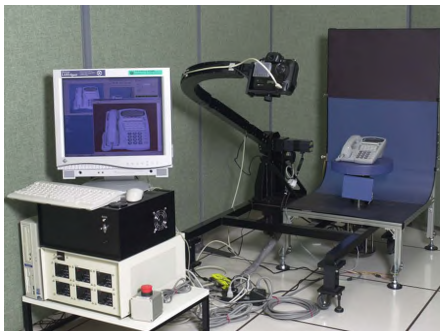


Fig. 1: The original 3D capturing system

\*e-mail: hmatsu@aecl.ntt.co.jp

### 2.2 The Sensor Cube

The sensor cube is the hook-up between the real world and the objects to be regenerated. It can also monitor environmental conditions around it. We have developed two kinds of sensor cubes, a geometry sensor and a light sensor (Fig. 2).

The geometry sensor has five distance sensors and five LEDs on each face and measures the distance to the surrounding real objects. The sensor detects “collisions” between virtual objects and real objects, and the LED turns red when a collision occurs. This allows users to arrange virtual objects in a room in the same way they would real objects.

The light sensor measures lighting conditions, such as the direction, intensity, and color of the light source. When a virtual object is rendered in real space, the information is used to decide the tone and shade of the surface of the object, the direction and intensity of its shadow, and so on, so that with the object fits naturally into the real world.

### 2.3 Viewer

We have developed a viewer (Fig. 3) and its software. The viewer consists of a tablet-type personal computer and a USB camera. We employ the ARToolKit [Schmalstieg et al. 2001] provided by the University of Washington for the registration of objects. Since views of the virtual objects are generated using the sensor information, harmonization between the real world and virtual objects is achieved (Fig. 4).

## 3 Conclusions

Regeneration of real objects is our ultimate goal, for which the system we propose is the first step. We believe that the realization of this concept will bring various attractive applications into our life.

## References

- MATSUOKA, H., TAKEUCHI, T., KITAZAWA, H., AND ONOZAWA, A. 2002. Representation of Pseudo Inter-reflection and Transparency by Considering Characteristics of Human Vision. *Proceedings of Eurographics 2002 (to be published)*
- Schmalstieg, D., Billinghurst, M., Azuma, R., Hollerer, T., Kato, H., and Poupyrev, I. 2001. Augmented Reality: The Interface is Everywhere. *Course Notes of SIGGRAPH 2001*



Fig. 2: The sensor cubes.



Fig. 3: The viewer.



Fig. 4: Example to show varying emotions depending on the view angle.



# Rich Curve Drawing

Suguru Saito  
Tokyo Inst. of Tech.  
P&I lab.

Akane Kani  
Tokyo Inst. of Tech.  
Faculty of Engineering

Youngha Chang Masayuki Nakajima  
Tokyo Inst. of Tech.  
Graduate School of Information  
Science and Engineering

## Abstract

In simple style drawings, like Comics and traditional cel animation, curved strokes are relatively important. Of course, the shape of the curve is the most important. However, subtle changes of curve width cannot be ignored. We propose a powerful method allowing subtle width changes to be applied to general 2D curve data. The algorithm is based on curvature information of the input curve, and keeps carefully the impression of the original curvature. The resulting image expresses a pen-and-ink drawing style.

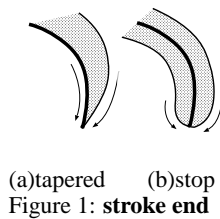
## 1 Introduction

Hand-drawn curves have variable width. The expression of the same curve is different when the width dynamics change. So variable width is an important element of curve expression.

We proposed a new method that applies a variable width according to a line's curvature. Therefore the algorithm potentially accepts a wide range of curve data: rendering curves from 3D models, curves input interactively, 2D edges extracted from images, and so on.

## 2 The goal of the method

This study was inspired from a study of the relationship between attractive curves and curvature[Nakajima et al. 2001]. So we focused on the curve's original curvature, and managed to keep the curvature in the produced outlines while varying the width. Here, we define the following constraints on the produced outline curve:



1. curves get fat asymmetrically; exaggerate the outside
2. the original line defines the center of the produced outlines, if it is straight
3. outline curves must keep their smoothness

Constraint 1 is necessary for curve ends which taper as a stroke reaches the end. If the symmetry is exaggerated, as in Figure 1(b), the radius of curvature of both outlines face different ways, and it expresses a stop. We aim to keep original curve's speed, therefore exaggeration of one side is required, as in Figure 1(a). Constraint 2 is necessary for correct expression of the original line position. Constraint 3 is necessary for natural reasons.

## 3 Algorithm

The algorithm of the curved part is based on an exaggeration method using the Inverse Geometric Heat Equation[Steiner et al. 1998]. When the input curve is  $C(s)$  and the partial curvature is  $\kappa(s)$ , the generated outline is  $C_{new}(s)$  in Equation (1).

$$C_{new}(s) = C(s) - f(\kappa(s))\hat{N}(s) \quad (1)$$

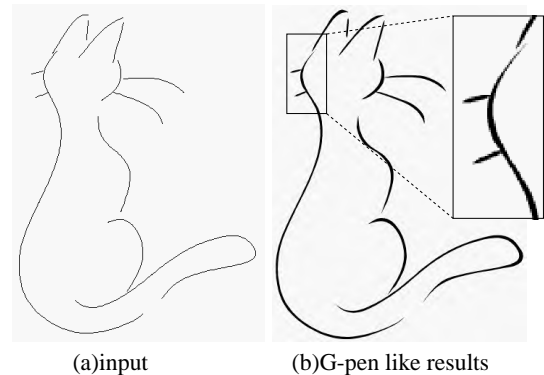
Here,  $\hat{N}(s)$  is the direction of the radius of partial curvature at  $s$ , and  $f$  is the exaggeration function depending on partial curvature.  $f$  becomes larger as  $\kappa$  becomes large.

At the curved part of the line, the exaggeration is on one side only. On the other hand, at the straight part, a new outline is generated on both sides of the line. The base width is predefined.

That is the basic process. In addition, we use two smoothing algorithms for outline connections at joints between curves and straight lines, and for concave points where curvature changes quickly and the exaggeration algorithm produces outlines whose curvature is incorrect.

## 4 Image results

We developed the algorithm for an interactive tool as an evaluation. The input curves in Figure 2(a) were drawn by a user with a pen tablet digitizer. The image has only shape information. Using this information, Figure 2(b) was produced automatically. Curves with variable width help achieve a more artistic drawing.



## 5 Conclusion and Future Work

We proposed an algorithm for drawing curves with variable width similar to G-pen style. By using only curvature information, the algorithm is applicable to various types of input sources. We plan to find better parameter sets in order to produce much better curves.

## References

- NAKAJIMA, N., KURIHARA, Y., HARADA, T., TAKAGI, S., AND YOSHIMOTO, F. 2001. Analysis of characteristics of curves in natural and artificial objects. *IPSJ SIGNotes 2001-CG-104 2001*, 89, 33–36.
- STEINER, A., KIMMEL, R., AND BRUCKSTEIN, A. M. 1998. Planar shape enhancement and exaggeration. *Graphical Models and Image Processing 60*, 2, 112–124.



# Sandwiching Surfaces

Jörg Peters\*, Xiaobin Wu

University of Florida

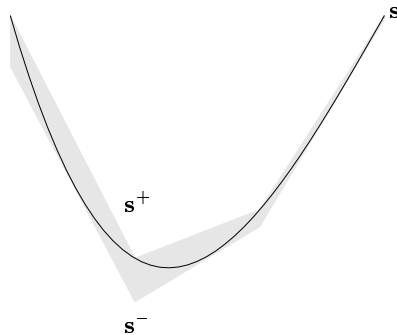


Figure 1: A slice through a three-piece sandwich.

Given a composite spline surface, we show how to efficiently construct two matching triangulations that sandwich the surface. Such a two-sided enclosure,  $(s^-, s^+)$ , supports collision detection, re-approximation for format conversion, meshing with tolerance, one-sided smoothing and silhouette detection as illustrated in Figure 2.

Measuring closeness to the silhouette or determining the distance between objects are fundamental issues in computer graphics. While efficient algorithms exist for piecewise linear surfaces with not too many pieces, objects in B-spline, Bézier or generalized subdivision representation pose numerical and implementation challenges due to the curved geometry. Naive linearization of a curved surface, say triangulation by sampling, incurs an error that is expensive (compared to sandwiches) to quantify and is particularly noticeable when estimating the range of the normal.

Sandwiching surfaces differ from existing bounding constructs by taking the geometry representation, say the Bernstein-Bézier form, into account and performing and tabulating certain pre-optimizations off-line, once and for all. This makes their computation particularly robust and cheap, eg for sandwiching a bicubic tensor-product patch, 8 tabulated floats, a linear combination of second differences and the comparison of a few determinants are required (the comparison is only needed to create continuity). The maximal error can be read off at the vertices.

Sandwiching surfaces differ from most existing bounding constructs by joining without gaps, i.e. yielding a continuous ‘upper’ and ‘lower’ triangulation.

Sandwiching surfaces are predictively refinable. While the width of enclosures converges quadratically in the worst case like most other bounding constructs, the constants dominating the width for the relevant first subdivisions are much smaller due to the pre-optimization.

Sandwiching surfaces are based on recent advances in enclosing polynomial functions [Lutterkort 2000; Lutterkort and Peters 2001a; Lutterkort and Peters 2001b]. To make these new techniques practically useful for computer graphics, function enclosures need to be generalized to *surfaces*. This extension is nontrivial and requires several modifications of the function approach; in particular, it gives rise to useful constructs called midpaths and midpatches that encode the surface.

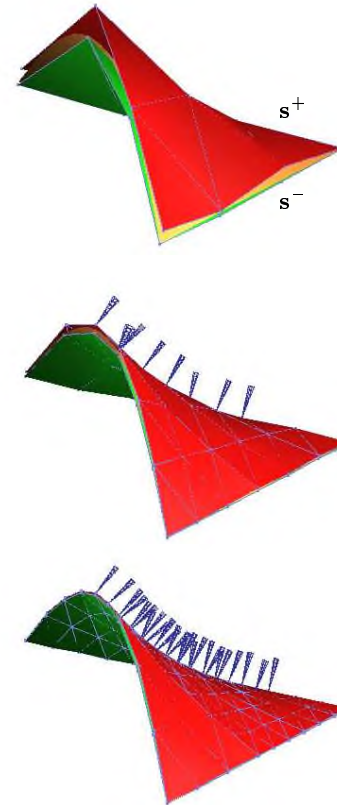


Figure 2: Adaptive refinement of geometry and normal sandwiches near the silhouette.

During the **presentation**, we will serve up fresh surface sandwiches in linear time, explain their ingredients, and compare them with fat arcs and spherical shells, convex hulls, k-dops, bounding circles and ellipses, axis-aligned and oriented bounding boxes. We will discuss robustness at singularities, how to rule out local self-intersections and demonstrate collision detection as well as refinement near silhouettes.

## References

- LUTTERKORT, D., AND PETERS, J. 2001. Optimized refinable enclosures of multivariate polynomial pieces. *Comput. Aided Geom. Design* 18, 9, 851–863.
- LUTTERKORT, D., AND PETERS, J. 2001. Tight linear bounds on the distance between a spline and its B-spline control polygon. *Numerische Mathematik* 89, 4, 735–784.
- LUTTERKORT, D. 2000. *Envelopes for Nonlinear Geometry*. PhD thesis, Purdue University.

\*NSF NYI CCR-9457806

# Shader Analytical Approximations for Terrain Animation in “The Time Machine”

Johnny Gibson  
Digital Domain

For *The Time Machine* the CG effects team faced the daunting task of producing imagery that represented long exposure and time-lapse photography of various types of terrain. The scope of the effort and the shot design largely preempted the use of simulated or exclusively explicit techniques for terrain construction. The challenge we faced was to procedurally model surface detail features into terrain that would visually represent the erosion of volumes of earth and rock over time. These features had to be consistent over time and through the entire volume of earth and rock through which the terrain would erode. These restrictions gave rise to two shader techniques used both in the procedural construction of geometry and in the displacement phase of terrain shading: bouldering and gulleying.

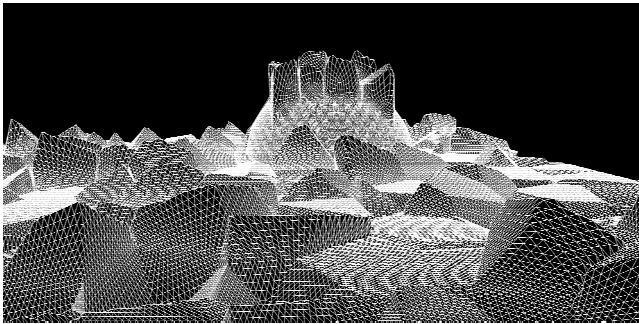


Figure 1. Bouldering

## Bouldering

The naive view of the task of filling a volume of sediment with various-sized boulder-like chunks of rock would dictate the population of the scene with pre-modeled rocks. Then, when the main body of terrain is animated through this space, it could pass through the rock models and expose them for a brief time.

There are three problems with this solution. The first is its inefficiency. The majority of rocks aren't in view at render time. The second is that this requires a surface intersection solution. The third is that there is no guarantee that the features derived from this system would inter-align the way boulders would. Viewers would like to see features that appear to have either been packed together or broken apart from other features. This requires a more cellular approach than the naive particle case.

The solution used by *The Time Machine* CG effects team was to modify the 3D Voronoi algorithm, abstracting it to a 3D density field that implicitly populated any terrain infinitely with boulder-features. We do not displace the geometry in a traditional way: taking its displacement magnitude from the source geometry's intersection with the Worley noise function. Instead, the source geometry is projected onto the surface of each goal feature. This has the advantage of forming deep cracks between features and multiple facets on their faces: in short, whole boulder-features.

Thus, when a source geometry animates through the field of Voronoi features, the features are consistent. They aren't relative to the source surface other than that it needs to be in a feature's domain for the feature to be exposed. When geometry is outside of the domain of any feature, its calculation is automatically skipped.

Finally, when bouldering is used as a shader solution, completely hidden features are never computed. The function is demonstrably boulder-like with cracks that form between features and multiple facets on the face of each feature.

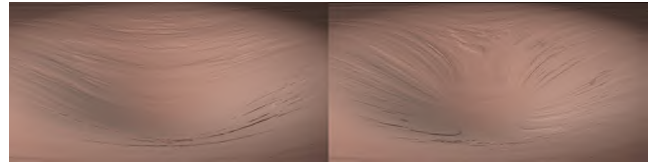


Figure 2. fBm with (left) low X frequency and (right) gulleying

## Gulleying

A natural assumption when designing shader displacement code for terrain is the use of fractal noise. A classic problem for the application of noise in a 3D volume, such as the above terrain case, is the alignment of noise to a surface. How does one align a low frequency axis of fractal noise to variously facing surfaces? For example, how does the shader-writer align noise to a mountain that has many 45-degree slopes facing various directions? At first glance there is no simple solution. Here, *The Time Machine* CG effects team presents a way to work around the gulleying problem.

First we pick an erosive flow direction for the source point being shaded. This is a solution that doesn't have to be complex because visual complexity will be contributed by a fractal lookup later. We choose the simplest procedural way we know. Flow direction is derived by projecting the normal of the surface at the source point onto the plane defined by the normal and the source point along the vertical axis. This projection gives a direction that is aligned to the surface at that point and generally faces down. This is the downward or erosive flow direction for this source point.

Once the downward flow direction is solved we need a way to derive a noise uvw coordinate from it. We use two ways to do this, both of which are valid but for different looks. The first is to project the source point along the flow vector onto a fixed horizontal plane. The second uses a large fixed sphere as the projection surface. The former has the effect of increasing the uvw frequency as the terrain becomes more flat. The latter reduces the effect of the frequency change between flatlands and vertical faces. The noise uvw coordinate is then the 3D spatial coordinate of the results of the above projection. Fractal noise functions work best for the film's needs but, since this solution only hinges on the texture lookup, any function can be used.

With this system, when noise is looked up, any source points near each other which also have similar erosive flow directions will have similar noise results. Any points that have the same erosive flow directions, which also lie along the same erosive flow path, will have exactly the same noise lookup. Any time the flow direction on a surface changes, the direction of the noise lookup changes thus affecting the frequency of the noise. This has the visual effect of long fractal features aligned to the faces of objects and with low frequencies applied along the downward (erosive flow) directions. We call them gullies.

# Shape Analogies

Aaron Hertzmann<sup>1</sup> Nuria Oliver<sup>2</sup> Brian Curless<sup>1</sup> Steven M. Seitz<sup>1</sup>

<sup>1</sup>University of Washington <sup>2</sup>Microsoft Research

## 1 Introduction.

This sketch presents “Shape Analogies,” a method for learning line styles from examples. With this approach, an artist or end-user simply draws in the desired style; the system analyzes the drawings and generates new imagery in the same style. For example, to design an outline style for a nervous character, one may draw a jittery stroke; to design an outline style for a robot, one may draw a very rigid style with many sharp angles.

Our approach, called Shape Analogies, builds upon recent work in image texture synthesis and transfer. Instead of generating pixel values, however, the goal is to generate a new curve from one or more input curves. This problem is challenging because we must somehow generate the output curve shape, its neighborhood parameterizations, and its correspondence to the source curve simultaneously. Moreover, the algorithm must account for variable rotations and translations between the drawings. Our approach is to iteratively reestimate curve shape, neighborhood sampling patterns and translational/rotational alignments.

## 2 Problem statement

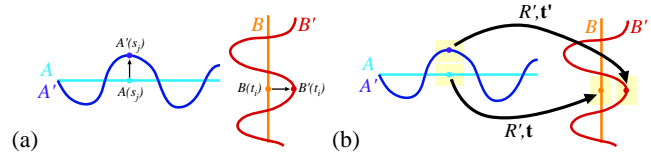
The shape analogies problem is as follows: given an example “unfiltered” curve  $A$  and example “filtered” curve  $A'$ , we would like to “learn” the transformation from  $A$  to  $A'$ , and apply it to a new curve  $B$  to generate an output curve  $B'$  (Figure 1). In short, we wish to generate  $B'$  to satisfy the relation  $A : A' :: B : B'$ .

Intuitively, the problem can be viewed as simultaneously optimizing for three goals: (1)  $B'$  should “look like”  $A'$ , (2) The *shape* of the  $B'$  curve should have the same relationship to the shape of  $B$  as the shape of  $A'$  has to  $A$ , and (3) The relative *positions and orientations* of the  $B'$  curve should have the same relationship to  $B$  as the positions and orientations of  $A'$  do to  $A$ . The first two terms directly correspond to the matching terms in the Image Analogies algorithms; the third term is new, and specific to shape modeling.

## 3 Algorithm

The simplest version of our algorithm is shape synthesis (Figure 2). After initializing the output curve with a short segment from the example, the output curve is extended incrementally. At each step, the end of the output curve is compared to an overlapping set of neighborhoods in the example curve. The output is then extended to “look like” the best-matching example neighborhood. Two neighborhoods are compared by aligning them (by Umeyama’s algorithm [1991]), and then summing the pairwise distance between pairs of corresponding points on the curves. Both of these steps use point correspondences determined by arc-length sampling from the respective neighborhoods. The alignment step ensures rotation and translation invariance in the drawing style.

The complete shape analogies algorithm generalizes this synthesis method by using the  $A$  and  $B$  curves to constrain the output (according to the problem statement above). Multiresolution synthesis allows efficient multipass refinement. We also support constraining the output curve to pass through specific points, in order to guarantee that the output drawing has a specific topology.



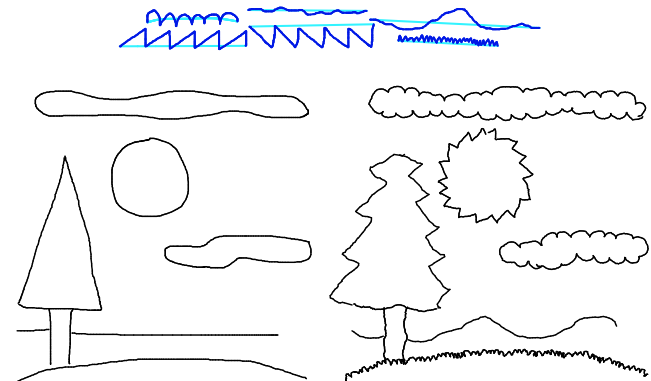
**Figure 1** Shape analogies problem statement. (a) Our goal is to generate a  $B'$  curve, such that, for every value  $t_i$ , there is some value  $s_j$  such that: the neighborhood around  $B'(t_i)$  looks like the neighborhood around  $A'(s_j)$ , the neighborhood around  $B(t_i)$  looks like the neighborhood around  $A(s_j)$ , and the offset between  $B$  and  $B'$  matches the offset between  $A$  and  $A'$ . (b) These comparisons are performed with respect to the rigid transformation that best matches the curve neighborhoods. (A different rigid transformation will be estimated for each candidate).



**Figure 2** Shape synthesis. The example shape is shown on the left, and the output shape on the right.



**Figure 3** Shape analogy applied to handwriting. The stylized “B” is generated by analogy to the other curves.



**Figure 4** Combining shape analogies. The curves in the left image were processed with learned styles and composited manually to produce the right image. The example  $A$  and  $A'$  curves are shown above.

## References

- EFROS, A. A., AND FREEMAN, W. T. 2001. Image Quilting for Texture Synthesis and Transfer. In *Proceedings of SIGGRAPH 2001*, 341–346.
- FREEMAN, W. T., TENENBAUM, J. B., AND PASZTOR, E. 1999. An example-based approach to style translation for line drawings. Tech. Rep. TR99-11, MERL, Feb.
- HARRISON, P. 2001. A Non-Hierarchical Procedure for Re-Synthesis of Complex Textures. *Proc. WCG*.
- HERTZMANN, A., JACOBS, C. E., OLIVER, N., CURLESS, B., AND SALESIN, D. H. 2001. Image Analogies. *Proceedings of SIGGRAPH 2001*, 327–340.
- UMEYAMA, S. 1991. Least-squares estimation of transformation parameters between two point patterns. *IEEE Trans. Pattern Anal. Machine Intell.* 13, 4, 376–380.

# Shape from Distortion: 3D Range Scanning of Mirroring Objects

Marco Tarini \*

Visual Computing Group, IEI, CNR, Pisa, Italy

Hendrik Lensch, Michael Goesele, Hans-Peter Seidel

Max-Planck-Institut für Informatik, Saarbrücken, Germany

Objects with mirroring surfaces are left out of the scope of most recent 3D scanning methods. We developed a new acquisition approach, shape-from-distortion, that focuses on that category of objects, requires only a still camera and a monitor, and generates high quality range scans (plus a normal field). Our contributions are a novel acquisition technique based on environment matting [Chuang et al. 2000] and corresponding geometry reconstruction method that recovers a very precise geometry model for mirroring objects.

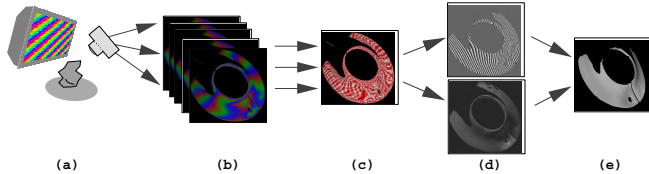


Figure 1: Acquisition Pipeline: Four images of reflected patterns are captured (a+b) and the matte (c) is extracted. A depth and a normal field are calculated (d) and combined to the final range scan (e).

## Data Acquisition

In order to obtain a range scan, we display a stripe patterns four times (each rotated by 45 degree resulting in different frequencies along the axes) on a monitor and capture the reflection of the monitor on the object's surface with a digital camera. From the captured images, we compute with subpixel accuracy which pixel on the monitor is reflected at each position.

The projected stripe pattern consists of linear ramps in the RGB color channels which add up to 1 at each position making the system robust to different albedo. Furthermore, appropriate calibration steps ensure that the whole optical system has a linear response. We can thus compute for a single captured image an offset value in the periodic stripe pattern that corresponds to each color triple. Given four images, we compute the exact monitor position that is reflected in each pixel using the interference properties of the patterns. The system is overdetermined and the redundancy is used to improve the accuracy of our approach.

In contrast to other spatio-temporal modulation approaches (e.g., [Rocchini et al. 2001]), this method does not rely on progressively lower frequency patterns and coherency between neighboring pixels so that even fast changing signals can be captured. The fact that the high frequency information from all captured images is used, increases the stability of the technique.

## Geometry Reconstruction

Given the captured matte, the internal camera parameters, and the position of the monitor relative to the camera, we can convert a normal directions into a depth values and vice versa by two methods: Using only local information of a single pixel, we can trace a ray trough each pixel on the image plane and reflect it of the surface so that it hits the corresponding pixel on the monitor (according to the matte). This constraint allows to directly calculate a depth value from a given normal and the other way around. Furthermore, given the normal and a depth value at a pixel and under the assumption of surface continuity depth values can be calculated for neighboring pixels by following the slope determined by the normal.

\*e-mail: mtarini@iei.pi.cnr.it

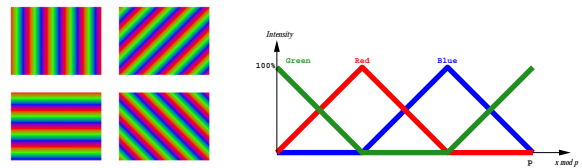


Figure 2: Projected Patterns: the patterns (left) and their basic structure.

We start with an arbitrary depth value at one pixel and calculate its corresponding normal. Then depth values in the neighborhood are computed following the slope determined by the normal; normals can then be determined at these points via the matte. This way we propagate an initial depth value through the whole field. The resulting depth and normal fields are then globally optimized following the slope along all directions (not only the direction of propagation).

Still, the depth of an initial point must be determined. To recover this value, we measure the (in)coherence of the resulting normal field i.e., the error made when a normal field is integrated along different paths. Incoherent normal fields are not justified by any depth field. In all test cases, the function relating initial depth values to the incoherence had a pronounced minimum corresponding to the correct depth value at the starting point. That minimum is quickly found using a golden section search.

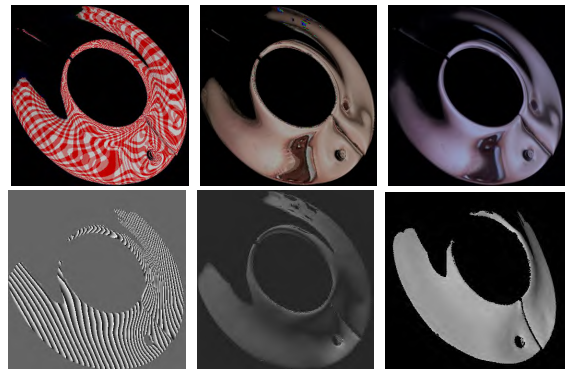


Figure 3: An Example: A matte captured over a plumbing object. Above: The images show the simulated reflection of checker pattern and an image of a face; the real reflection of the same image is on the right. Below: the resulting depth map (isovalue lines) and normal map calculates from the matte, and the final 3D range scan.

## Results

We successfully applied our technique to acquire a number of mirroring objects. As can be seen in Figure 3, even tiny features of the objects can be captured well. To quantify the accuracy of the system, we scanned a mirroring sphere of 60 mm diameter resulting in an error of less than 0.02 mm.

## References

- CHUANG, Y. ET AL. Environment matting extensions: Towards higher accuracy and Real-Time capture. In *SIGGRAPH 2000*, 121–130.
- ROCCHINI, C., CIGNONI, P., MONTANI, C., PINGI, P., AND SCOPIGNO, R. A low cost optical 3D scanner based on structured light. In *EG 2001*, 299–308.



# Shape-Based Character Animation

Eric Guaglione  
Walt Disney Feature Animation  
[eric.guaglione@disney.com](mailto:eric.guaglione@disney.com)

Doug Sweetland  
Pixar  
[doug@pixar.com](mailto:doug@pixar.com)

In character animation, it is often said that a strong pose is supported by its' silhouette. Since the shape of a 3D character represented in a graphical, 2D screen-space is ultimately our goal, shouldn't our tools better support posing the silhouette? Yet commercial software packages have focused on implementations to pose the character from the inside out, rather than looking at the shape inward.

At Walt Disney Feature Animation, we have the additional challenge of realizing 3D versions of beloved traditionally drawn characters. Since everyone knows how Roger Rabbit or Genie from "Aladdin" are supposed to look, the viewer adopts a more critical eye. Characters such as these have taught us how to refine our tools, but we feel CGI tools are still lacking in controlling shape-based character animation.

- How can one use existing tools in a different fashion to support the notion of shape-based animation?
- What problems still arise from available tools?
- What are the pitfalls of approaching 3D character animation with a shape-based paradigm?
- What tools are still lacking to control a character's exterior shape?

Pixar cleverly designs characters to support the style and appeal of animation they have become so well known for. Typically Pixar's characters are not as amorphous as "Aladdin's" Genie, but Randall from "Monsters, Inc." is a classic example of an original character whose shape defines his pose. Pixar will discuss their approaches to shape-based animation relative to their unique style of character design.

This sketch proposes to demonstrate some of the shape-based character animation approaches Walt Disney Feature Animation and Pixar have taken in the past few years, while raising questions for the future.



standard deformation

shape controlled





# Sketchy Rendering

John Haddon  
The Moving Picture Company  
theboyhaddon@hotmail.com

## Abstract

The animated short "a flatpack project" was intended to emulate the aesthetics of traditional art techniques in a digital medium. This necessitated the creation of custom rendering and image processing code to reproduce the appearance of line drawings in both pencil and ink, along with other effects such as the bleeding of ink in water and the application of pastel to paper. This sketch describes some of the techniques used in achieving these ends.

## Overview

The bulk of the techniques used focused on the 2d post processing of images rendered from 3d scenes modelled and animated in Maya. The post process approach was largely motivated by a need to "colour outside the lines", something not easily achievable in a conventional renderer. Initial output typically comprised of depth, facing ratio and tonal information encoded in an RGB image. These intermediate images were then processed variously to provide the elements of the final animation.

## Line Drawings

The first stage in the creation of line drawings was the familiar cel line generation process. Edge detection was performed on the facing ratio and depth images, yielding clean line images which required further processing to achieve the sketched look. The lines were then traced to extract a curve definition. Rather than use any of the more complex curve fitting techniques, regularly spaced samples along the lines were used to form the control points of an interpolating cardinal spline. Although less precise than more rigorous curve fitting techniques, this was adequate (and perhaps even more appropriate) for the sketchy look desired, with the tension properties of the cardinal spline proving invaluable later in the processing.

Once extracted, the cardinal splines were plotted using different styles to emulate the effect of both pencil and ink. It was found that by plotting each spline over a (very) widely varying range of tensions and combining the results, a good semblance of a pencil sketch was achievable. Higher tensioned plots tended to give a loose line, much like an initial sketch. This was then filled in with lower tensioned plots tightening to a more refined line. By controlling the range of tensions a variety of looks were achieved, with very high tensions causing straight line segments to overshoot wildly, yielding the appearance of construction lines.

In contrast, inked lines were typically plotted at a single, more sensible tension. Line width was influenced by both position along the stroke and the tonal value beneath the brush, enabling shadow areas to be more thickly delineated.

## Ink bleeding

Rather than use a more complex simulated method, a very simple cellular automata rule was developed to provide the effect of ink on a wet page. The resulting greyscale images of bled ink were

enhanced by remapping through color gradients, giving the appearance of the separation of the ink. Although not having all the richness of a more complex approach (e.g [Curtis et al. 1997]), these techniques proved adequate for the job in hand.

## Pastel on paper

The effect of pastel on paper was one of the more easily achieved effects. Tonal images were roughened with a paper texture, and then remapped through a gradient composed of blocks of solid colour, using a simple expression node in Chalice. Mapping the midtone transparency allows the paper to show through, preventing the page from becoming saturated with heavy colour.

## Compositing

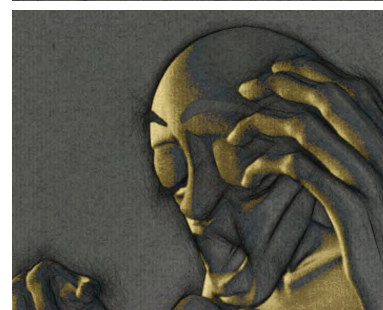
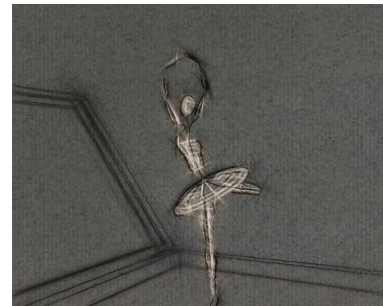
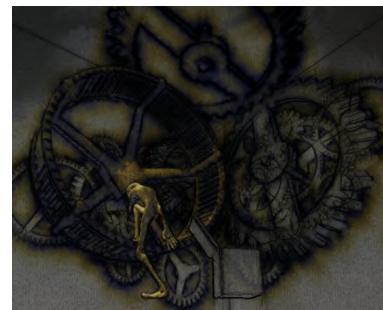
As the pastel example shows, a fair amount can be achieved through the use of simple compositing tricks. The flexibility and immediacy of the 2d application are believed to provide the best environment for considering issues of style and aesthetics. For this reason the technical processes described were used to generate a great many elements, each of which provided added flexibility in the compositing of the final image.

## Conclusion

It is hoped that this sketch has shown that a combination of relatively simple computer graphics techniques can be employed along with a little thought, creativity and good fortune to produce pleasing images in the vein of traditional media.

## References

CURTIS C., ANDERSON S., SEIMS J., FLEISCHER, K. AND SALESIN, D. 1997. Computer-Generated Watercolor. In *Proceedings of ACM SIGGRAPH 1997*.



# Slow Growing Volumetric Subdivision

Valerio Pascucci\*

CASC – Lawrence Livermore National Laboratory, CA 94551

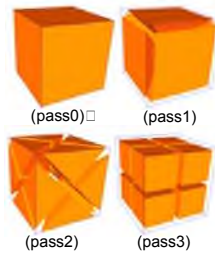
## 1 Introduction

In recent years subdivision methods have been successfully applied to the multi-resolution representation and compression of surface meshes. Unfortunately their use in the volumetric case has remained impractical because of the use of tensor-product generalizations that induce an excessive growth of the mesh size before sufficient number is preformed. This technical sketch presents a new subdivision technique that refines volumetric (and higher-dimensional) meshes at the same rate of surface meshes. The scheme builds adaptive refinements of a mesh without using special decompositions of the cells connecting different levels of resolution. Lower dimensional “sharp” features are also handled directly in a natural way. The averaging rules allow to reproduce the same smoothness of the two best known previous tensor-product refinement methods [Bajaj et al. 2001; MacCracken and Joy 1996].

## 2 Slow Growing Volumetric Subdivision

The volumetric subdivision is a generalization of the 4-8 surface subdivision scheme introduced recently in [Velho and Zorin 2001].

In the 3D case we organize each level of subdivision in four passes, from 0 to 3, where pass 3 of level  $l$  is coincident with pass 0 of level  $l+1$  as shown in the figure on the right.

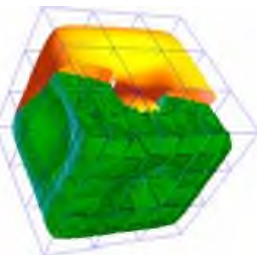


**From pass 0 to pass 1.** For each cell  $c$  of the input mesh, a new vertex  $p$  is inserted at its center. The cell  $c$  is partitioned in pyramids by connecting  $p$  to each facet of  $c$ . Pairs of pyramids with the same base facet  $f$  are merged together into a cell  $F$ .

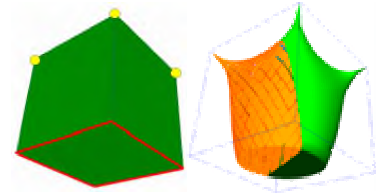
**From pass 1 to pass 2.** For each facet  $f$  of the base mesh, a new vertex  $q$  is inserted at its center. The cell  $F$  containing  $f$  is divided in pyramids by connecting  $q$  to each facet of  $F$ . Groups of pyramids sharing a base edge  $e$  are merged into a cell  $E$ .

**From pass 2 to pass 3.** For each edge  $e$  of the base mesh, a new vertex  $r$  is inserted at its center. The cell  $E$  containing  $e$  is partitioned in pyramids by connecting  $r$  to each facet of  $E$ . Groups of pyramids contained in the same base cell  $c$  and incident to the base vertex  $v$  are merged into a single cell.

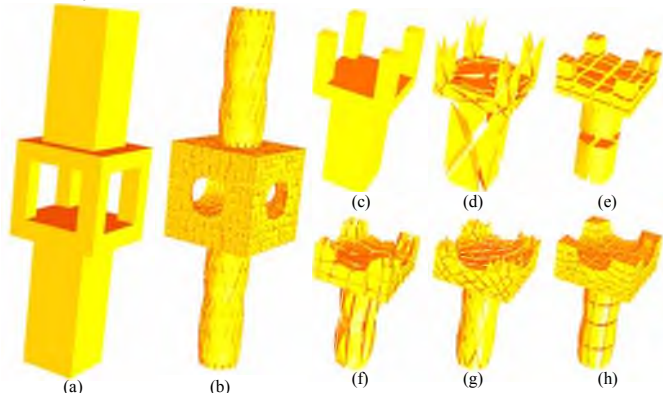
This subdivision in four passes refines gradually a mesh by decomposing first its cells, then its facets and finally its edges. A unified solution for boundary cases, for sharp features and for adaptive refinements, is obtained simply by skipping the merging stages. For example if  $f$  is a boundary facet there is only one pyramid with base  $f$  (the cell  $F$  will have  $f$  on its boundary). Similarly, if one refines only one of two cells adjacent on  $f$ , there is only one pyramid with base  $f$  and the resulting mesh has no hanging node. Therefore, adaptive refinements (see figure on the right) do not require special



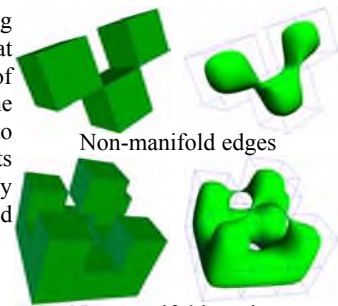
treatment. The figure on the right shows the refinement of a cube, where the four top vertices (yellow) are four sharp 0-features, while the bottom edges (red) form a sharp 1-feature. Note that the base mesh is not required to have hexahedral cells. For example, cells like the dodecahedron below are subdivided directly.



The figure below shows the mesh constructed for a mechanical piece. (a) is the base mesh and (b) the mesh after six refinements. (c-h) are the intermediate steps of the refinement (bottom part of the object). Note that a tensor-product scheme would introduce the same number of vertices in just two refinements, requiring usually more vertices to achieve the same smoothness.



In conclusion the slow growing subdivision is a general tool that allows the practical use of subdivision methods in the volumetric case, extends to higher dimensions and admits base meshes that are virtually unrestricted (e.g. non-mifold features as in the right figure).



## Bibliography

- Bajaj, C.L., Warren, J., AND Xu, G. 2001. A smooth subdivision scheme for hexaedral meshes. Submitted to “The Visual Computer”. TR-01-05, UT at Austin, April.
- MacCracken R., AND Joy, K.I. 1996. Free-form deformations with lattices of arbitrary topology. In *Computer Graphics (SIGGRAPH'96)*, pages 181--188.
- Velho, L., AND Zorin, D. 2001. 4-8 subdivision. *Computer Aided Geometric Design*, volume 18, Issue 5, pp 397-427.

\* [pascucci@llnl.gov](mailto:pascucci@llnl.gov). This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract no. W-7405-Eng-48



# Star Fields in 2D

Maria Giannakouros  
Digital Domain \*

## Introduction

Night scenes filmed on indoor stage sets, and in day or evening light, all have one thing in common—the absence of a starry sky. Even shots filmed at night, under perfect conditions, fail to capture stars. Since the focus is on local subjects which need adequate lighting to be captured, the sky is too under exposed to capture stars.

Techniques for adding stars to night scenes have ranged from glitter on black cloth to 3D renders of spheres. Creating or making changes to a background plate using the previously mentioned techniques can be time consuming and at this at a very early stage of production. Regardless of the method for generating stars, once the background star plates are generated, the final stage is to composite the background with the foreground or action scene. An efficient solution would be a tool that adds stars to the composite at this time.

## Criteria for Realistic Stars

The criteria for a successful alternative method were speed, appearance, and adaptability. First, the new method had to be faster than the existing 3d render solution. Second, the new stars had to look as good as or better than the 3d rendered stars. Third, the method had to be adaptable or at least have the potential to be adaptable. Adaptability, in this cases implied having the ability to fulfill production requests. Some initial production requests were: adding stars to static scenes; gradually adding stars to a scene such as simulating a change in the camera focal length; and simulating movement through space.

## Speed

Speeding up the process was done by using a very simple render and anti-alias algorithm. The goal was to draw the smallest stars possible and still give the impression of small dots rather than square pixels. Small anti-aliased squares look like rendered spheres on film. Using this simple render, the same data set of 9,000 stars could be rendered at 9 seconds/frame versus over 2 minutes/frame for the 3d render. An added benefit of the anti-aliased squares is the star maintains a hot center which helped preserve the intensity of the star when motion blurred.

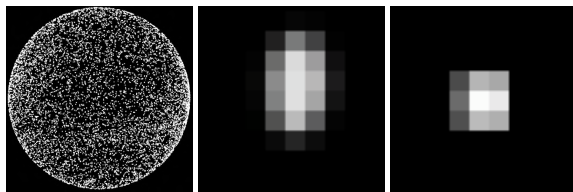


Figure 1: 3d data set of 9,000 stars(l), 3d render results(c), Star field render(r)

## Appearance

The goal for the appearance of the stars was to generate the smallest possible points of light that moved with consistent intensity across the screen. Stars with 1x1 area worked well as long as the rgb values of the star did not exceed one. Stars with this small area appeared to blink as the center of the star traveled across the anti-aliased region

(Figure 2). However, High Dynamic Range values are especially useful when trying to preserve star intensity when motion blurred.

The problem of blinking stars was corrected by increasing the area of stars with rgb values greater than one. All bright stars were drawn as 2x2 squares. The bright center was now distributed over a larger region so it was less apparent when the star center traveled across the anti-aliased region.

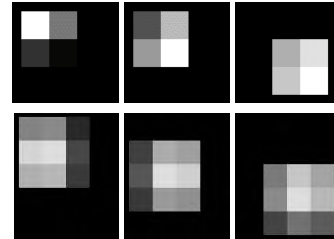


Figure 2: Progression of a bright star center traveling across a 2x2(t) and 3x3(b) anti-aliased region

## Adaptability

The final goal of this project was to come up with a design that was adaptable to future requests. Some of the initial requests were: to achieve realistic motion blur, and to show a night sky as the camera focal length changed.

The goal for motion blur was to maintain the star intensity as the camera moved. This was done by using HDR values for the rgb values of the stars. The star values were clipped when the camera was at rest. As the camera moved, the HDR values were averaged over the blur region to produce an average intensity that was near one. Thus the motion blurred stars were still visibly bright.

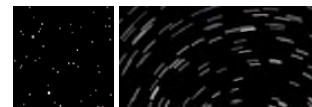


Figure 3: Stars without motion blur(l) and Stars with motion blur(r)

The goal for camera focal length changes was to adjust the star count according to the camera's focal length. As the focal length decreases, more dim stars should become visible and bright stars should maintain their brightness. This is currently done by animating the brightness value, as the camera focal length decreases, the brightness values of the stars increase.

## Future Work

Camera focal length changes could also be solved by accounting for the z-values of the stars. Currently, data is read from a file that provides a z-value so that all the stars are projected on a sphere (Figure 1). However, z-depth information could be added and brightness could be made a factor of distance. Then, as the camera focal length changes, distant stars come into view. This solves the problem of simulation changes in camera focal length more accurately. Now that distance and brightness information are dependent on one another, simulating travel through star fields is also addressed.

\*mgianna@d2.com



# "Still I Rise" Painterly Animation off the Shelf

Umesh Shukla  
5809 Abernathy Drive  
Los Angeles AC 90045  
shukla@umeshshukla.com

## Abstract

Painterly animation has till now been a part of research groups and enthusiast. This film attempts to bring it closer to the masses using commercially available software.

**Keywords:** Painterly renderings

## 1 Introduction

During a 1984 Siggraph panel discussion on the History of Computer Graphics, Alan Kay, often credited as the father of personal computers, identified the main challenge for computer graphics to be “to move from the concrete to the abstract, from pretty pictures that render the invisible concrete to computer graphics that have meaning in everyday life, from images that are there simply to be recognized to symbols that are there to be understood”.

I have always felt motivated by this comment. These motivations have led me to explore painterly renderings in a variety of forms. Painterly animations have long been a part of the palette of a traditional animator. The traditional animator has used paint on glass, color pencils, pin screen and even plasticine on glass amongst a wide variety of techniques to create beautiful images. The film Still I Rise attempts to explore painterly looks using common commercially available software. The film uses a combination of 2d and 3d packages and a mix of techniques to create a dream sequence.

## 2 Inspiration

The film takes its base on the coincidental suicides of the Elephant Man Joseph Merrick and impressionist painter Vincent Van Gogh. Tired of sleeping in a seated position, with his head on his knees, Joseph Merrick took a mortal risk and lay flat on his bed. This suffocated him and ended his life. The incident took place in April 1890.

During the month of July in the same year, an as yet unknown master of impressionist art took his own life. His name was Vincent Van Gogh.

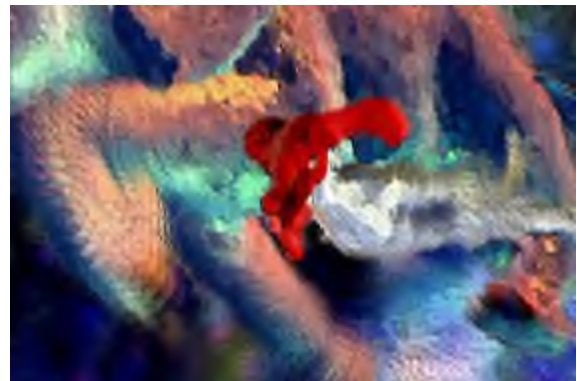
The thought that these two unique individuals lived at the same time and embraced their own deaths became the basis of inspiration for the film. The film depicts Joseph's last dream, as he breathes his last, and visualized as an impressionist canvas.



## 2.1 Details

The focus of the exercise was to achieve a painterly look as the final result. Initial tests with various methods and packages were conducted. Poser was chosen as the animation package primarily because of its vast library of rigged models and their easy availability. Poser also fulfilled any animation need the short had with great ease. Models were either bought from online vendors, or were acquired through the Poser news group where individuals sell their creations. Models were altered to suit the need of the short, as and when it was needed.

Initial lighting was done in Poser and rendered layers were created for each element present in the frame. Photoshop was used to create texture maps required for the models. Rendering was done in layers to keep a finer control on the final colors. These rendered layers were passed through Re:visionFX Video Gogh plug-in to generate the painterly look. Numerous such layers were created.



These layers were taken in to Houdini for color correction. Additional effects were added in Houdini to achieve the desired look. The composited layers were edited together with the music in Adobe Premiere.

Final renders were transferred to CD's and passed on to an outside facility for printing on film.

## 3 Conclusions

I am sure it is the beginning of a new era. As more and more artists begin to use computers to create, various new styles of animation will emerge. Vincent Van Gogh, in letters to his brother Theo, once wrote, “The painter of the future will be a colorist such as has never yet existed. I believe in the absolute necessity of a new art of color, of design, and – of the artistic life. And if we work in that faith, it seems to me that there is a chance that we do not hope in vain.”

I am sure we do not.

# Stylized Flowing Hair Controlled With NURBS Surfaces

Ramon Montoya-Vozmediano  
Walt Disney Feature Animation  
Ramon.Montoya@disney.com

Mark Hammel  
Walt Disney Feature Animation  
Mark.Hammel@disney.com

A project currently in production at Walt Disney Feature Animation requires realistic hair with a high degree of artistic control. In this sketch we present a set of techniques that achieve these goals along the production pipeline.

## Modeling Hair

Look development artists traditionally groom hair or fur as a final step, prior to rendering. Because of this, modelers or animators do not have direct control over final shapes and silhouettes. We would like to be able to use standard tools to model the hair shapes. A solution to this problem, specially suited for long hairs, is to model a volume using NURBS to define the shape of the hair.

We fill the volume with *instanced* hairs, represented with a NURBS curve and a set of properties per control point. Each instanced hair corresponds to the linear interpolation of as many *key* hairs as possible, so that the hair does not respond locally but as a cohesive whole to the deformation of the surface. Initial key hairs are grown on the surface's isocurves and also inside the enclosed volume. All this is done at render time, the artists just work on the original NURBS surface.

## Animation and Dynamics

The generated hair can be animated using standard techniques on the original surface. The key-framed animation drives the movement, but for increased realism and detail we add a layer of secondary dynamics. This particular work is underwater, and a basic forward dynamics explicit solver was found to be adequate. Hairs are modeled as a set of linear and angular springs.

A relaxation process combines the original animation and the dynamics, its intensity controlled by a factor set by the user. This factor can be supplied through a texture map painted on the NURBS surface, so that the roots follow the animation and the tips have a more dynamic movement.

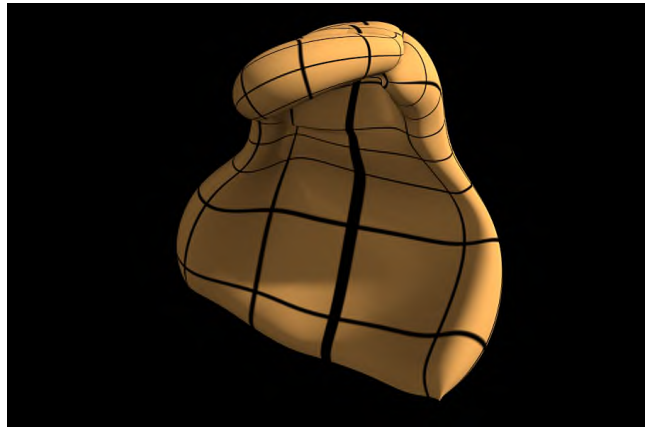
To modify the way the simulation affects instanced hairs, we treat the resulting dynamics as offsets associated with control points on the key hairs. Later in the pipeline, look development can decide if whole locks or just isolated hairs will deviate from the original surface, and also by what amount. This is a vital control to follow the animated shapes and provide realistic visual cues at the same time.

## Shading and Efficient Rendering

We used a RenderMan DSO to implement the hair growth. Hair shaders have access to the weighted sums of properties derived from the original surface, like (u,v) coordinates. Averaged surface normals are used for lighting. This results in a more stylized appearance and standard lighting techniques, usually not valid on hair, can be used.

At render time, the instanced hairs are generated and grouped inside extruded triangles that run along the length of the shape. We further decompose these tubes by splitting the hair curves into their piecewise Bezier form. This provides an intermediate layer of small bounding boxes so that at any one time only the curve segments in a few boxes are instanced into memory. If other hairs or a scene element occlude a box, RenderMan just discards it. This gives us automatic culling and reduces the memory restriction for high numbers of hairs.

*Special thanks to Marcus Hobbs and Stephen Poehlein.*



Original NURBS surface



Final rendered hair

# Synchronous Pronouncement

Sandra Villarreal  
Pratt Institute  
sandra@villarrealstudio.com

## Abstract

Synchronous Pronouncement is a generative interactive installation that will explore the aesthetics of an interactive medium as an extension of our body. The irregularity and unpredictability of our world is a result of unique patterns left behind by movements. The installation will be a representation of interacting processes that operate in nature. Autonomous behaviors generated by code will be triggered by the presence of users for the creation of a unique real-time audio-visual experience.

## 1 Introduction

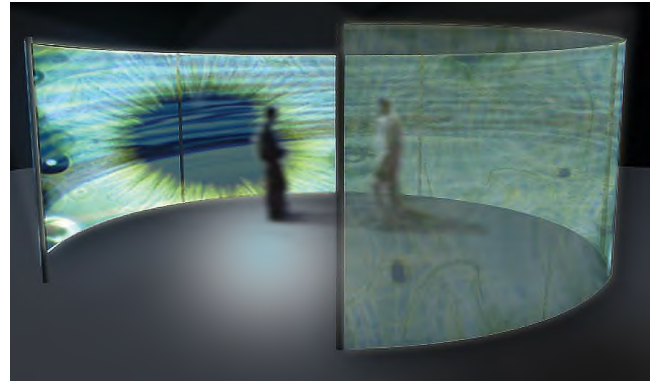
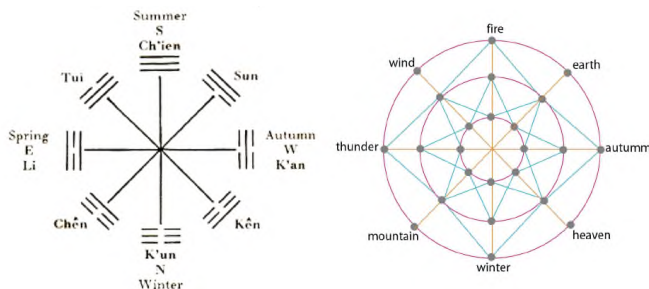
The word Pronouncement evokes an oracular process that will be hidden in my artwork. The users will be walking in an environment whose basic hidden structure will be related to the hexagrams presented in The Book of Changes (I Ching). The lines of the hexagrams change and regroup themselves, just as world situations continually change and recombine themselves. The hexagrams provide complete images of conditions and relationships existing in the world. In the same manner, the movement of the user from one area to another will generate different associations and combinations of imagery and sound unique to the specific interaction.

## 2 Exposition

Synchronous Pronouncement will use synthetic audio and projections of animated abstract imagery displayed on a 300 degree semi-transparent wall. Each person entering the area will activate a sensor and depending on his position, an image will fill the screen. The image will be related to the element of the sensitive area that the user will activate. As the visitor moves through the area, new elements will appear and evolve, depending on the path followed by the user. The eight sensors correspond to fire, earth, lakes, heaven, water, mountains, thunder and wind. Each sensor's emission cone will be divided in zones defined by its sensing ranges.

## 3 Imagery

The images will be abstract but will reveal, through closer observation, structures that will resemble the appearance of forms in nature. Cellular structures will evolve from the dark as if the viewer was immersing in a microscopic view or a telescopic view of a galaxy. Different types of structures will be able to melt, disperse or collapse. Sounds will evolve with the images and will accentuate the sense of spatiality in the installation. They will basically be synthetic sounds resembling sounds from nature.

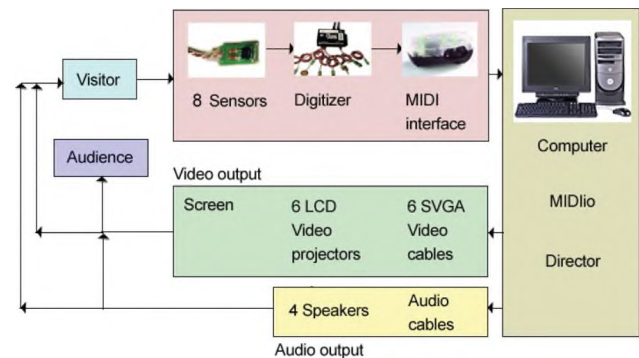


## 4 Technology

Eight sensors (FarReach I-CubeX from Infusion Systems) will transmit information to the computer about the position, movement patterns, and the number of users in the area. The computer will act as a central controller collecting the input information through a digitizer that converts the analog signals generated by the sensors into MIDI signals. The digitizer will be connected to a MIDI controller (I-CubeX by Infusion Systems). It works with software applications such as Max as well as 3rd-party software plugins such as MIDIio for Macromedia Director. MIDIio provides MIDI input and output functions to enable Lingo control and responses to externally generated MIDI messages. An interactive application built with Macromedia Director will output the imagery. The imagery will be developed using Imaging Lingo with Macromedia Director. Generative algorithms will generate autonomous behaviors.

## 5 Conclusions and Future Work

The evolution of new interactive devices has primarily focused on two of our senses: sight and hearing. Another important sense is touch. The tactile world creates a sense of place and movement. It is impossible to completely interact with the world while being attached to a fixed device because it creates a contradiction between the dynamic content and the fixed physical interaction. Our sense of touch is also linked with other senses, such as weight, body, posture and balance. Interactive media should evolve as extensions of our bodies permitting a flow and a transparent continuous dialogue between the users and the devices. Interactive communications should be based on natural experiences and responses. Those are aspects I intend to explore through the development of this project and future related projects.





# Tangible Viewpoints: A Physical Interface for Exploring Character-Driven Narratives

Ali Mazalek, Glorianna Davenport, Hiroshi Ishii  
MIT Media Laboratory  
{mazalek, gid, ishii}@media.mit.edu

## 1 Introduction

Over the centuries, stories have moved from the physical environment (around campfires and on the stage), to the printed page, then to movie, television, and computer screens. Today, using wireless and tag sensing technologies, story creators are able to bring digital stories back into our physical environment. The Tangible Viewpoints project explores how physical objects and augmented surfaces can be used as tangible embodiments of different character perspectives in a multiple point-of-view interactive narrative. These graspable surrogates provide a more direct mode of navigation to the story world, bringing us closer to bridging the gap between the separate realms of bits and atoms within the field of digital storytelling.

## 2 The Interface

The Tangible Viewpoints interface uses wireless graspable pawns to navigate through a multiple viewpoint story. Each pawn represents a particular character in the story, and when placed on the sensing surface, the story segments associated with its character's point-of-view are projected around it. A small lens-like object serves as a selection tool for viewing the story content. By gliding this selection tool over particular story segments, users can playback their associated story content on a nearby monitor. This causes the narrative to advance and new segments to become available. An aura is projected around each pawn to give a visual representation of the prominence of that character viewpoint in the current telling of the story. Changes in the story space (i.e. when the story moves forward or there is a shift in character perspective) are thus reflected by dynamic changes in the projected graphics.

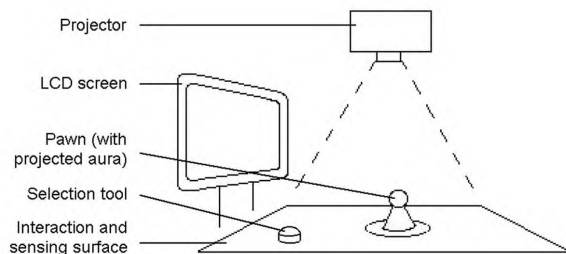


Figure 1. Diagram of the Tangible Viewpoints setup.

## 3 Story Content and Narrative Model

The Tangible Viewpoints narrative structure supports character-driven stories, in which different characters with distinct personalities and unique points-of-view are used as the primary basis for creating the narrative. The different segments of a multiple point-of-view story are organized according to the character viewpoint they represent, as well as their place in the



Figure 2. A user viewing a story in Tangible Viewpoints.

overall narrative. As the narrative unfolds, the system gathers information about which characters a user has been interacting with, and makes decisions about what segments to present based on this knowledge. If the user chooses to focus their interest on a particular character, the system narrows the story scope as it progresses, resulting in greater depth in that character's story. On the other hand, if the user spreads their focus between the three characters, the result is a much broader story.

When two pawns are touched together on the interaction surface, the system displays only the portions of story that are relevant to both characters. In this way, users can examine relationships between the characters and see how their individual stories or perspectives relate to one another.

The story segments in Tangible Viewpoints can consist of various types of media (video, audio, images, text), and as such can present character development, action and location with as much complexity as any scene of a film or chapter of a book.

## 4 Future Extensions

We have used Tangible Viewpoints for text and image-based stories (such as the example shown in Figure 2), and are now exploring the creation of an acoustic story space consisting of character dialogues, sound effects and ambient background sounds, coupled with the projection of a visual landscape onto the interaction surface. In this case, the visuals serve to give users an idea of the layout of the story's physical location. The story audio can then be mapped to different parts of the stereo field based on each pawn's position and movements on the interaction surface, creating a "soundscape" for the story that enhances the projected visual landscape.

## References

MAZALEK, A. 2001. *Tangible Interfaces for Interactive Point-of-View Narratives*. MS thesis, MIT Media Laboratory.



# Techniques for Interactive Audience Participation

**Dan Maynes-Aminzade**  
MIT Media Lab  
monzy@mit.edu

**Randy Pausch**  
Carnegie Mellon University  
pausch@cs.cmu.edu

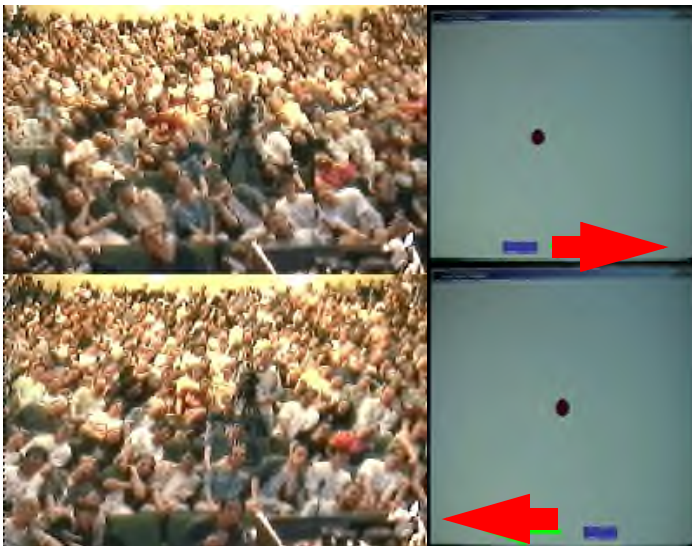
**Steve Seitz**  
University of Washington  
seitz@cs.washington.edu

## INTRODUCTION

At SIGGRAPH in 1991, Loren and Rachel Carpenter unveiled an interactive entertainment system that allowed members of a large audience to control an onscreen game using red and green reflective paddles. In the spirit of this approach, we present a new set of techniques that enable members of an audience to participate, either cooperatively or competitively, in shared entertainment experiences. Our techniques allow audiences with hundreds of people to control onscreen activity by (1) leaning left and right in their seats, (2) batting a beach ball while its shadow is used as a pointing device, and (3) pointing laser pointers at the screen. All of these techniques can be implemented with inexpensive, off the shelf hardware.

### 1. AUDIENCE MOVEMENT TRACKING

Audience movement tracking allows audience members to control an onscreen game without the use of physical props. We position a camera at the front of the auditorium and point it at the audience. We then encourage the members of the audience to move in certain ways, and analyze the streaming video to control an onscreen game. For example, audience members can lean left or right in their chairs to steer a race car or move a paddle in the video game *Pong*.



**Figure 1:** A large audience leans left and right to control an onscreen game of *Pong*.

### 2. BEACH BALL SHADOWS

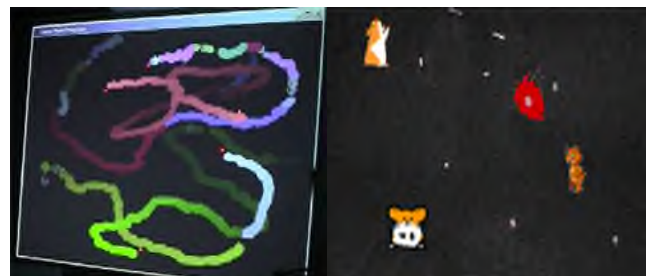
The common crowd activity of batting a beach ball before a concert provided the inspiration for our second interaction technique. We project a game on the front screen of a movie theater, and as the audience bats a beach ball into the air, the ball casts a shadow on the screen. We point a camera at the screen and use computer vision techniques to track the ball's shadow. This tracking allows the audience to play interactive games using the shadow of the ball as a cursor.



**Figure 2:** Missile Command using beach ball shadow tracking. The beach ball (top left) casts a shadow (bottom right) which acts as a cursor.

### 3. LASER POINTER TRACKING

It is not uncommon for members of a large movie theater audience to shine laser pointers at the movie screen before the film begins. By pointing a camera at the screen and tracking the dots, we can create compelling interactive entertainment experiences. A sample application of our system, shown in Figure 3, is a collaborative paint program that allows the audience to draw graffiti on the screen. Another example is a game that encourages rapid flocking of laser points to particular locations. Figure 3 shows one such game, a laser version of the carnival midway game "Whack-a-Mole." In this game, the audience members use their laser pointers to hit a sequence of targets. We built other games where the audience mixed audio samples like a disc jockey, engaged in live polls and trivia competitions, uncovered hidden images, and played an interactive multi-user version of connect-the-dots.



**Figure 3:** A laser pointer paint program (left) and laser Whack-a-Mole (right).

## REFERENCES

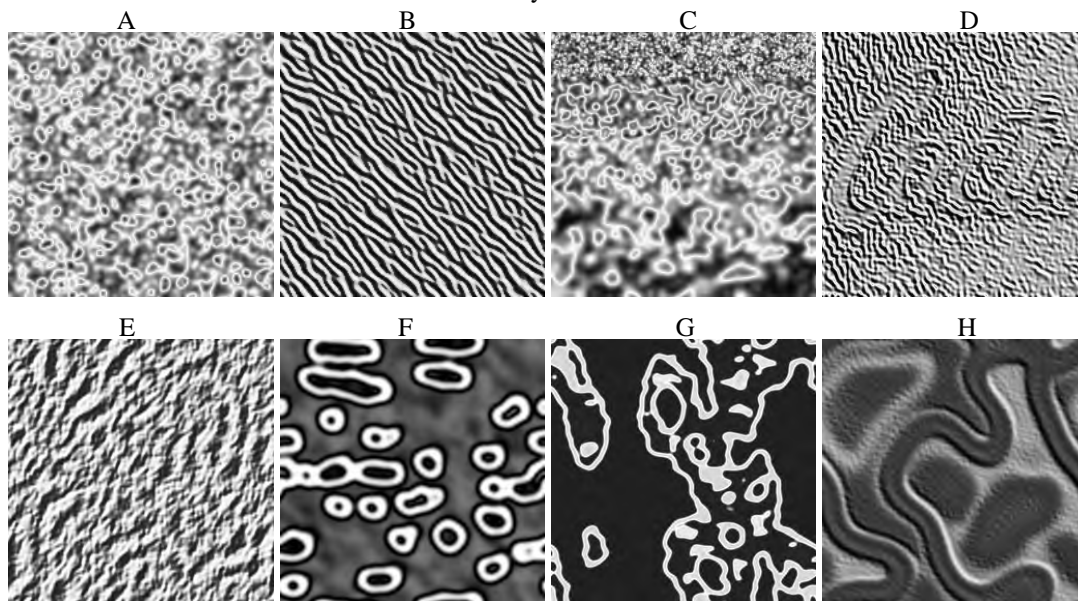
1. CARPENTER, L., Cinematrix, Video Imaging Method and Apparatus for Audience Participation. US Patents #5210604 (1993) and #5365266 (1994).

# Textures from Nonlinear Dynamical Cascades

David Mould\*

Eugene Fiume†

University of Toronto



## Introduction

A great deal of attention has been paid to the problem of texture synthesis. Procedural techniques have become commonplace. Yet while a proliferation of models, and the success of nonparametric synthesis-from-example methods, has made an extraordinary variety of textures realizable, the creation of novel textures remains a challenge.

We propose a model that combines iteration of nonlinear functions and convolution with multiscale cascade. The textures resulting from this model share characteristics with textures we deem “natural”: they exhibit a tension between structure, or smoothness, and randomness, or roughness.

## The Model

Our method makes use of iterated nonlinear maps [Devaney 1989; Ott 1993] which provide roughness. Organization is provided to us by the convolution. While repeated convolution yields only trivial textures, and repeated iteration of pointwise maps produces a garble of uncorrelated values, interleaving nonlinear maps with convolution steps can prevent either eventuality. Consider the following schema:

$$A_{i+1} = p(N(A_i)).$$

In the above,  $p$  is a map and  $N$  is a convolution; both act on a lattice  $A$ . The subscript denotes iteration, a discrete dynamical process.

This equation describes a texture synthesis system. Populating a lattice with uniform noise and repeatedly applying a convolution and a map produces structured texture; we stop after a fixed number of iterations (say 5) and use the transient patterns then present. In the interest of speed, we use a small convolution kernel, no larger than  $7 \times 7$ . Textures (A) and (B) were generated by this method.

Because the convolution is linear, it is closed under linear transforms such as scales and rotations. In (C) above, we show an example of a texture where the kernel has been dilated by different

amounts in different regions of the image. We may also apply the operator to structured initial conditions rather than to noise. Features in the original image are then expressed in texture features. We show an example of this in (D).

An operator acting on a structured image preserves some large-scale features of that image; we can exploit this in a cascade process wherein lower-frequency levels are upsampled and used as initial conditions for higher-frequency levels. The final texture is the last level of the cascade. Larger-scale details remain, embellished by the action of later operators. Images (E)–(H) above are examples of textures from the cascade. The basic system generates textures with frequency content on a single scale; the cascade admits textures with content on arbitrarily many scales, a characteristic often identified as “natural”.

## Summary

We present a system for texture synthesis geared towards the computation of natural textures. Natural textures, in our perception, are those characterized by the interplay between smoothing and roughening. Our system is able to rapidly synthesize texture from uniform noise. Given structured input, the system can be used for texture transfer. Multiscale output can be achieved through a cascade of operators. Despite the simple formulation, the system is capable of producing a complex array of texture images.

## References

- DEVANEY, R. 1989. *Introduction to Chaotic Dynamical Systems*. Addison-Wesley, Redwood City.
- OTT, E. 1993. *Chaos in Dynamical Systems*. Cambridge University Press, Melbourne.
- TURK, G. 2001. Texture synthesis on surfaces. In *Computer Graphics (SIGGRAPH '01 Proceedings)*, 347–354.
- WITKIN, A., AND KASS, M. 1991. Reaction-diffusion textures. *Computer Graphics (Proceedings of SIGGRAPH 91)* 25, 4 (July), 299–308.

\*mould@dgp.utoronto.ca

†elf@dgp.utoronto.ca

# The AcceleGlove: A Whole-Hand Input Device for Virtual Reality

Jose L. Hernandez-Rebollar, Nicholas Kyriakopoulos and Robert W. Lindeman  
The George Washington University  
[jreboll | kyriak | gogo]@seas.gwu.edu

## Abstract.

We present The AcceleGlove, a novel whole-hand input device to manipulate three different virtual objects: a virtual hand, icons on a virtual desktop and a virtual keyboard using the 26 postures of the American Sign Language (ASL) alphabet.

## 1 Introduction.

The AcceleGlove is a set of six dual-axis accelerometers, mounted on the fingers and the back of the palm, that report position with respect to the gravitational vector. Sensors are placed on the back of the middle phalanges, on the back of the distal phalange of the thumb, and on the back of the palm [figure 1]. To determine postures using this arrangement, we tested the 26 letters of the ASL alphabet until ambiguity was eliminated.



Figure 1  
The AcceleGlove

The Y-axis points towards the fingertip, and the X-axis measures hand roll or individual finger abduction when finger is in the vertical position.

Gloves like the CyberGlove and DataGlove have the disadvantage that readings change depending on how well it fits user's hand then, hand size is considered a source of noise. Since the AcceleGlove is not a complete glove, but a set of rings, hand size is not a source of error; hand thickness may affect interpretation of hand shape in some extreme cases, though. For example, when closing the fist, thin fingers are able to bend the joints up to 90° but thick fingers are not, so an 'a' in ASL could be miss classified as an 'o'.

## 2 Applications.

The CyberGlove, the DataGlove, the Dexterous Hand Master (DHM), and the PowerGlove, report only hand shape (finger flexion) [Sturman 1992]. The AcceleGlove provides hand posture, composed of hand shape and hand orientation as a 12-byte vector, at 9600 bits per second. Next, we present three virtual objects to show potential applications of the AcceleGlove.

## 2.1 Virtual Fingers

In general, only the thumb is able to move independently of the other fingers; the ability is gradually lost from the index to the pinkie. To track individual finger flexion, the AcceleGlove is connected to drive a virtual hand we have developed. The algorithm maps finger position into virtual finger bending. Figure 2 shows a lateral view of a virtual hand with fingers partially bent.

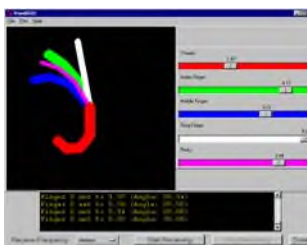


Figure 2 Virtual Hand

## 2.2 Grasp-and-Drop

To give a more intuitive interface between the user and a desktop PC, the AcceleGlove is used as a substitute for the mouse. The micro controller maps two hand shapes (closed fist or 'grasp' and open hand or 'drop') and Pitch and Roll of the hand into button 'clicks' and velocity control using the Microsoft-mouse format, so no especial hardware or driver is required to perform click-drag-drop operations.

## 2.3 Virtual ASL Keyboard.

Our third and more-complex application is a virtual keyboard controlled by ASL alphabet postures. It is used to practice or to learn the finger spelling technique. The user performs the hand posture wearing the AcceleGlove and, after our hierarchical, multi-class posture recognition algorithm solves for the letter, the key on the virtual keyboard is activated and the picture of the corresponding letter is shown.

The algorithm extracts a set of features that represents a posture (hand shape, palm orientation) unambiguously in "posture space." Hand shape is represented by overall finger flexion and overall finger roll. Palm orientation is given by pitch of the sixth sensor. With these three features, patterns are visualized as clusters in a 3-dimensional posture space. To our knowledge, this is a novel approach to ASL alphabet recognition. The GUI is shown in figure 3 with letter 'L' displayed at the left.

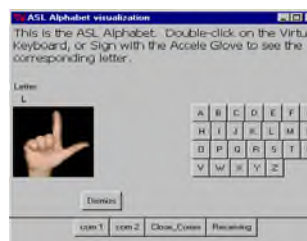


Figure 3 ASL Keyboard

## References

STURMAN, D. J. 1992. *Whole-Hand Input*. Ph.D. Thesis, Media Lab, MIT, Cambridge, Mass.



# The AR-ENIGMA – A PDA based Interactive Illustration

Volker Paelke, Joerg Stoecklein, Lennart Groetzbach, Christian Geiger, Christian Reimann, Waldemar Rosenbach  
University of Paderborn, C-LAB, Fuerstenallee 11, 33102 Paderborn, Germany  
{ vox, ozone, robak, chris, reimann, bobka }@c-lab.de

The AR-ENIGMA combines a PDA (personal digital assistant) with a camera, a high-speed wireless network connection and AR (augmented reality) technology to enable museum visitors to interact with an Enigma encryption machine.

## 1. The ENIGMA

The Enigma is a mechanical encryption machine that was used by the German military in WW2. It became famous when – after decades of secrecy – the successful attacks of the Allied codebreakers became public knowledge in the 1970s. Since then the Enigma continues to captivate public interest due to the vital role that intelligence from deciphered Enigma intercepts played in the Allied war effort and because the seemingly impossible task of breaking the Enigma ciphers was achieved by a fascinating combination of mathematical brilliance, daring espionage actions and one of the first “industrialised” information processing applications: the electro-mechanical “bomb” designed by Alan Turing and Gordon Welchman that is one of the predecessors of today’s computers.

## 2. The AR-ENIGMA Illustration

The AR-ENIGMA is an interactive illustration developed in collaboration with the Heinz Nixdorf MuseumsForum [HNF] that enables museum visitors to actively explore and experience the operation of an Enigma encryption machine. The illustration provides background information on a wide range of topics ranging from the Enigma’s historical context, over the basics of cryptography to the mathematical foundations of the enciphering algorithm implemented by the Enigma and the techniques employed by the successful Allied codebreakers.

Visitors use a modified PDA with a back-mounted video camera for interaction and to view the illustration. Since it is not viable to have numerous users interact with the actual exhibit an AR approach for interaction and information presentation was implemented: Markers captured by the camera are used to derive the user’s position and orientation relative to the physical exhibit. Based on this information a context sensitive presentation is generated that augments the actual camera view with additional information, e.g to explain the function and operation of different components of the Enigma, including those hidden within the machine like the substitution rotors. Illustration techniques like semi-transparency and explosion views

are used to provide visual access to machine details like the internal wiring of the substitution rotors.

In an interactive simulation mode visitors can use a virtual keyboard on the PDA display to encipher and decipher messages. The events that correspond to this enciphering in a real Enigma machine (e.g. key-presses, rotor movements) are illustrated in a 3D model of the Enigma that is registered with it’s real-world counterpart and displayed on the PDA. The simulation allows to configure arbitrary start-states (selection and position of rotors that correspond to different keys in the Enigma’s cipher) and enables several visitors to explore the functionality of the Enigma independently while the physical exhibit remains safely in its glass box.

The development of the AR-Enigma is an ongoing project. The next step will be the extension of the user interface to enable dynamic configuration of the stecker-board (a second encryption element in addition to the rotors) that is currently only supported within the simulation and a usability evaluation of different interaction and presentation techniques with actual museum visitors.

## 3. The Technology

The intended use of the AR-Enigma illustration by large audiences in a museum context precludes the use of bulky, complex and expensive hardware. The AR-ENIGMA design is therefore tailored for small hand-held devices based on standard hardware like Pocket PC PDAs, IEEE802.11b wireless LAN and mini video cameras. The current prototype is implemented on a Compaq iPaq PDA with a Lucent WLAN card and a mini video camera with a wireless up-link to a PC server. The PC server handles the computationally intensive tasks of image based object registration and 3D graphics generation that are beyond the processing power of current PDAs. Image based object registration was implemented using the AR-Toolkit [AR-Toolkit] while the AR-ENIGMA application logic and 3D presentation generation were implemented in Tcl and C++ using Open GL for rendering. The resulting video-stream is transmitted from the PC to the PDA through the WLAN connection that also transmits user interactions to the PC server.

## 4. References

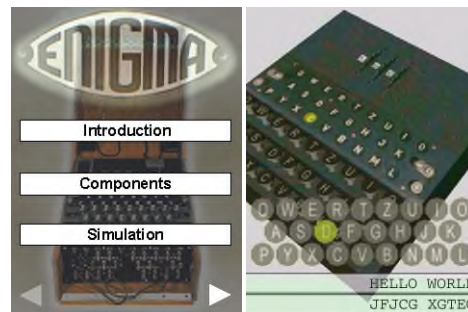
- [AR-Toolkit] [www.hitl.washington.edu/research/shared\\_space/download/](http://www.hitl.washington.edu/research/shared_space/download/)
- [HNF] [www.hnf.de/index\\_en.html](http://www.hnf.de/index_en.html)



Original Enigma in the Heinz Nixdorf MuseumsForum



Test Setup: PDA with Camera, WLAN card, video up-link and Enigma Mock-up



Interactive Enigma Simulation in AR-ENIGMA



# The Development of a Functional Visualization System for the Creation of Digital Human Models

John Jay Miller, Weidong Wang, Gavin Jenkins  
Mississippi State University  
jmiller@sarc.msstate.edu

## Motivation

Inadequately or poorly designed environments and tools of daily living impose barriers to people with a disability. This issue needs to be addressed in order for people with disabilities to lead full and purposeful lives. To accomplish this goal it is imperative that designers of environments and artifacts have an in-depth knowledge of human functioning in the performance of tasks and problem-solving strategies to develop environments and products that best accommodate performance of these tasks. They require an understanding and useful characterizations of the abilities of people with disabilities and relevant mechanisms to incorporate this into a modern design process.

## Goal

This project is an attempt to realize a Functional Visualization System for the creation of Digital Human Models (DHM) of people with disabilities for use by designers and clinicians. The System includes standard measurement procedures for the capture of anthropometric data, the development of computer programs and processes for the visualization of the data which has been captured, the verification of collected data and utilization of visualization tools by designers.

## Approach

The project uses a holistic approach for the generation of solutions to typical problems confronted by people with disabilities. This project is, by its very nature, multidisciplinary and broad in scope. It will provide services and information to many groups in order to demonstrate an integrated approach to disabilities studies and provide a proof of concept demonstration that products and environments, which meet the needs of people with disabilities, can move out the realm of 'orphan technology'.

Stakeholders involved in this effort include the following: Occupational Therapist/Clinicians, Computer Programmers, Visual Communication Designers, Industrial Designers, Statistician, Participants/End User / Specifier (people with disabilities) of products created using the visualization software.

## Specific Objectives

This effort to improve and utilize DHM technology for design purposes is important to designers and people with disabilities for two key reasons:

- Lack of Designer-Oriented DHM software, and the subsequent,
- Lack of Ergonomically Correct Products and Environments which are useful to people with disabilities.

## Preliminary Data

Preliminary work done by these researchers has demonstrated the ability to create useful DHM visualizations of individuals with disabilities, based upon simple measurements of limb segment length and maximum joint angles. [Miller and Wang 2001 ; Miller et al. 2000a 2000b]

Creating such useful models helps to overcome the specific shortcomings of traditional anthropometric visualization

approaches. In particular, static approximations of human motion which cannot:

- address the dynamic nature of functional abilities
- accommodate widely varying movement capabilities and different levels of total body engagement
- dynamically incorporate anthropometric factors into the design process

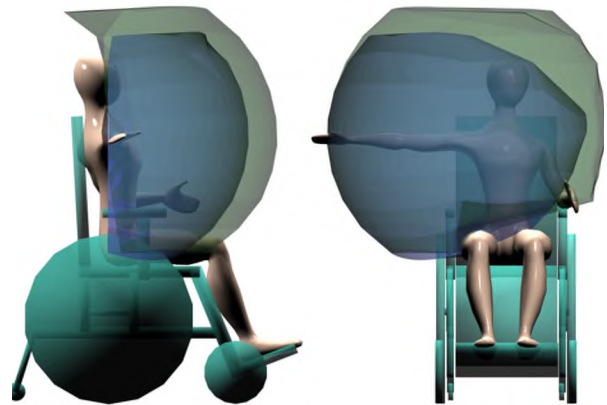
Using Commercial Off The Shelf (COTS) software the researchers have extended the basic software's capabilities to create 'data which makes data'. Rather than replaying choreographed animations, or repetitively positioning a static mannequin, designers can measure and manipulate three-dimensional objects which represent Range of Motion (ROM) volumes for specific individuals or 'percentile' ranges.

See also:

[www.sarc.msstate.edu/gstudents/wwang/web/research.html](http://www.sarc.msstate.edu/gstudents/wwang/web/research.html)



ROM surface generation sequence showing ROM (blue) intersection with cabinetry.



ROM Comparison for 41 yr. old female with arthritis (smaller blue/dark volume) and a 95th percentile female normalized for scale

## References

- Miller, John J and Wang Weidong. 2001. Application of Computer Visualization Techniques for the Use of Anthropometric Data in the Design Process, In *Proceedings of the International Conference on Affective Human Factors Design*, ASEAN Academic Press LTD, England UK
- Miller, John J, Wang, W., Jenkins G. 2000a The Development of Computer-based Representations of Anthropometric and Biomechanical Data for Use in Environmental and Product Design, Poster at *Infodesign 2000*, Coventry England
- Miller, John J, Wang, W., Jenkins G. 2000b The Anthropometric Measurement and Modeling Project, In *Proceedings of ACADIA2000*, Washington DC

# The Free-form Light Stage

Vincent Masselus\*, Philip Dutré and Frederik Anrys  
Department of Computer Science, K.U.Leuven

The Free-form Light Stage captures the reflectance field of an object using a freely movable, hand-held light source. By photographing the object under different illumination conditions, we are able to render the object under any lighting condition, using a linear combination of basis images. Our technique builds on recent techniques, such as the Light Stage [Debevec et al. 2000], where light sources are placed at fixed and known positions (e.g. using a gantry). We remove this limitation, and put no restrictions on light source placement.

Our system setup is shown in Figure 1. Photographs from a fixed camera position are recorded continuously, while a hand-held, calibrated light source is moved around the object. This process takes about 25 minutes, and typically a few hundred images are recorded.

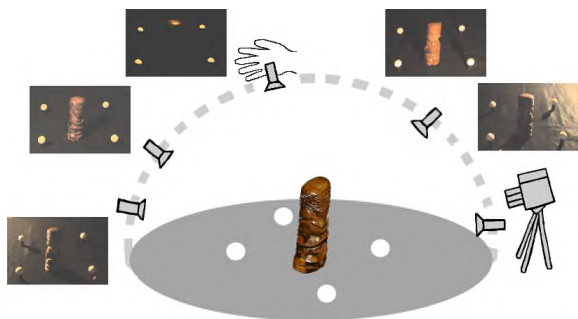


Figure 1: A small statuette is illuminated by a hand-held light source and continuously photographed, resulting in several pictures with different illumination.

The position of the light source in each picture is unknown. The lighting direction is estimated by placing small diffuse spheres next to the object. From the shading gradient on these spheres, the light source direction can be extracted [Lee and Rosenfeld 1985]. Care has to be taken that the presence of the spheres does not influence the illumination of the object. We found this not to be a significant problem in practice.

Once the light source direction is determined for each picture, the object can be relit using an arbitrary light map. All original images are linearly combined, their respective weights being determined by the location of their corresponding light direction in the light map. To account for the non-uniform distribution of light source directions, all light source directions extracted from the photographs are plotted in a concentric hemisphere map, and a Voronoi diagram is constructed. The weight of each photograph in the linear combination is determined by integrating the intensities of the light map over the appropriate Voronoi cell. Thus, we have re-rendered the object lit by the new light map (Figure 2). The Voronoi diagram is an elegant solution to the non-uniform distribution of light source directions over the hemisphere, which results from the uncontrolled and free movement of the hand-held light source.

The non-local reflectance field of almost any object or scene can be captured. The size and positioning of the object is limited only by the ability of a person to move the light source around the object. Objects fixed to their surroundings, such as mosaic floors or archaeological findings are in our reach, as are outdoor objects such as plants, small trees, or statues. However, since light maps represent the light incident at a single point, the object cannot become

too large, otherwise the assumption that the object can be faithfully rendered under the new light map becomes invalid.

Relighting the object can be accomplished easily, using real or artificial environment maps. Thus, real objects can be placed in virtual or real scenes.

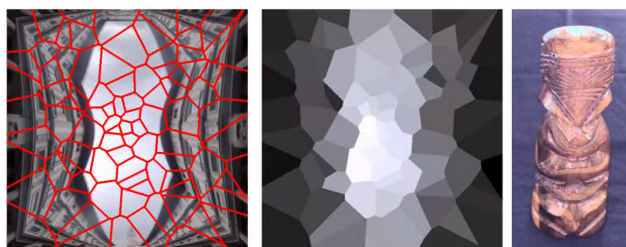


Figure 2: The concentric light map of the Uffizi gallery with super-imposed Voronoi diagram (<http://www.debevec.org/Probes>); the weight for each Voronoi cell; and the re-rendered object.

Future work includes a more accurate integration of the light map, experimenting with large objects and scenes, and try to develop a system where both light source and camera can be moved. Such a system would combine advantages of both the Light Stage and the Lumigraph, and would provide new views of any object, under any lighting condition.

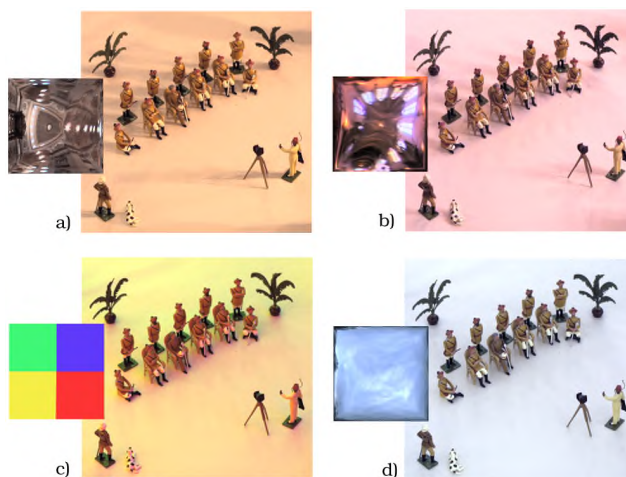


Figure 3: A portrait of toy soldiers rendered from 500 basis images. Note the reproduction of soft shadows (b,d), glossy highlights (a,on shoulders), and the use of an artificial light map (c).

## References

- DEBEVEC, P., HAWKINS, T., TCHOU, C., DUIKER, H.-P., SAROKIN, W., AND SAGAR, M. 2000. Acquiring the reflectance field of a human face. In *Proceedings of SIGGRAPH 2000*, K. Akeley, Ed., Computer Graphics Proceedings, Annual Conference Series, ACM Siggraph, 145–156.
- LEE, C. H., AND ROSENFELD, A. 1985. Improved methods of estimating shape from shading using the light source coordinate system. *Artificial Intelligence* 26, 125–143.

\*e-mail: {vincentm,phil,frederik}@cs.kuleuven.ac.be

# The Meditation Chamber : A Debriefing

A. Fleming Seay  
HCII  
Carnegie Mellon University  
afseay@cs.cmu.edu

Diane Gromala, Larry Hodges, Chris Shaw  
GVU Center  
Georgia Institute of Technology  
{gromala, hodges, cdshaw}@cc.gatech.edu

## Introduction

During the Emerging Technologies exhibition at Siggraph 2001, over 400 attendees experienced The Meditation Chamber. This immersive, bio-interactive environment was designed to use visual, audio, and tactile cues to create, guide, and maintain a user's guided relaxation and meditation experience. During this sketch, the project's producers will discuss the design and implementation of this unique installation. We will also show footage from the experience and discuss the subjective relaxation measures and the GSR, heart rate and respiration data generated by 411 Siggraph attendees.

## Design and Instrumentation

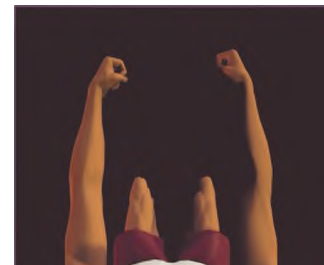
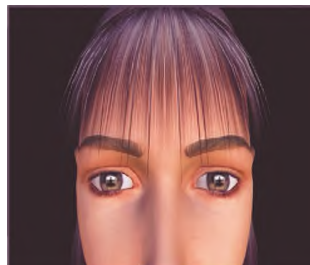
The system was designed over a period of several months by an interdisciplinary team of computer scientists, psychologists, artists, and educators. The end product was a three phase experience that offered a sampling of basic meditation and relaxation techniques to a user wearing a head-mounted display and a collection of biometric sensors. The three phases included an initial relaxation phase, a progressive muscle relaxation phase, and closed with an extended meditation called "following your breath". Galvanic skin response, respiration, and blood volume pulse were collected using the ProComp+, a commercially available device produced by Thought Technologies. The real-time biometric data was used to partially control the environment, shaping the user's experience.

## Subjective and Objective Results

A t-test showed that self report, post-session relaxation ratings ( $M=8.00$ ,  $SD=1.69$ ) were significantly higher than pre-session ratings ( $M=5.63$ ,  $SD=1.75$ ),  $t(410)=-24.45$ ,  $p=.0001$ . A ten-point Likert-type scale was used. Forty-nine of the 411 participants reported equal or lower levels of relaxation following the experience, with 18 of these reporting the same level, and 30 reporting decreases in relaxation of 1-3 scale points. The other 362 participants reported levels of relaxation 1-8 points higher following the experience, with a mean difference of 2.88 scale points between pre and post session ratings. There were no gender differences on either relaxation rating.

With respect to objective measures, preliminary analyses show that two general patterns of GSR profile can account for nearly 75% of the generated data and are each generally accompanied by two distinct respiration profiles. Just over half of the participants exhibit what can be called a novice GSR profile. This means that their GSR level starts relatively high, descends through the first phase of the experience, picks back up and shows peaks in the muscle relaxation phase, and then begins to decline again in the final phase, ending up at or beneath the low established in the first phase. Breathing patterns in individuals exhibiting the novice profile tend to be steadier and deeper in the final phase than in the first phase.

The second profile, which accounts for nearly a quarter of the GSR data, is termed the expert profile. Individuals exhibiting this profile show precipitous drops in GSR during the first phase, entering a very low and often flat GSR state before the muscle relaxation phase begins. This flat-line state is typically maintained throughout the remaining two phases of the experience, and is accompanied by a very steady but not necessarily deep breathing pattern. Individuals exhibiting the expert GSR profile also show very consistent respiration rate and amplitude throughout the experience. As we move forward, we plan to employ Principal Component Analysis to verify these preliminary findings.



Images clockwise from top: 1) A sunset image from the first phase of the experience where the user's GSR controlled the speed of the sunset. 2) A first person perspective of the male arm-flex exercise from the progressive muscle relaxation phase. 3) A third person perspective of the female eye-flex exercise from the progressive muscle relaxation phase.



# The Nail-Mounted Tactile Display for the behavior modeling

Hideyuki ANDO

Japan Science and Technology  
Corporation (JST)  
smartfinger@star.t.u-tokyo.ac.jp

Takeshi MIKI

The University of Tokyo

Masahiko INAMI

The University of Tokyo

Taro MAEDA

The University of Tokyo / JST

The "SmartFinger" is a new type of tactile display, which is worn on the nail side of the finger. It does not inhibit our tactile sensation, since the ball of the finger is naked and we can feel the environment directly. It is important to insert nothing between a finger and an object.

The "SmartFinger" can provide extra tactile sensations. For example, someone wearing this small device on his or her nail can touch a drawing on a flat sheet of paper and feel a virtual edge along the drawing's line, and virtual roughness from its texture pattern. This is a wearable device to add information of bump mapping on real surface objects using existing printing technique.

The "SmartFinger" consists of three components:

(1) Measuring the changing environment in real-time with photo-detectors [1]. This consists of a light emitting diode and a photo sensor, and makes the tip small by the optical fiber. For the invisibility of information, we used the ultraviolet light.

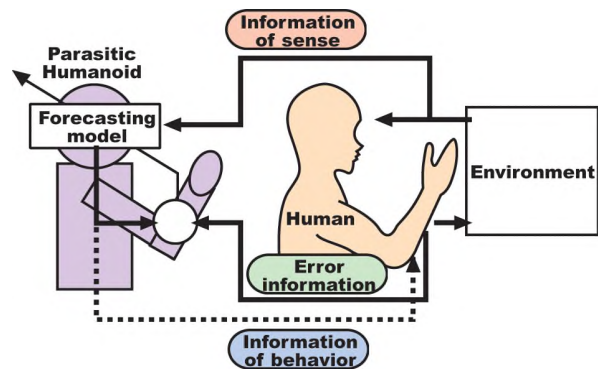
(2) Measuring change in the fingernail's color from the nail side to detect external force [2]. This consists of a light emitting diode and some photo sensors. These measure the amount of blood flows of the fingertip.

(3) Generating vibration on the nail to realize various tactile sensations. We provide tactile sensory output using the actuator. In order to give vibration limited in the vertical direction of the finger to this actuator, not a small vibration motor but the voice coil was used.

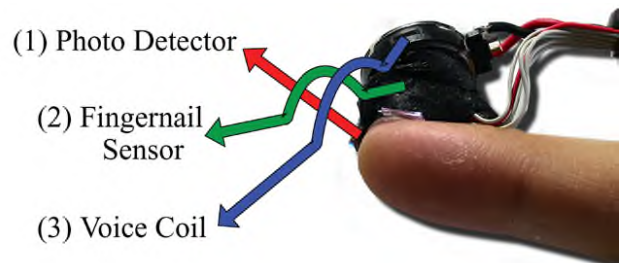
We research The Parasitic Humanoid (PH) that is a wearable robot for modeling of nonverbal human behavior. A user of PH should be able to obtain data with analyzed result without suffering any inconvenience in daily life. To realize such comfortably wearable PH, we need to develop measurement devices, which don't interfere with human behavior. We think that SmartFinger is suitable for this device.

Therefore, since visually impaired people make extensive use of their tactile sense. For example, they can use touch panel using the device detecting information from computer display. Because, the "Smart Finger" has a sensor equivalent to human's feeling input, and the actuator, which returns the stimulus to human. Therefore, it is possible to use as an interface that has a symbiotic relationship with human. For example, it made to notice what is not understood in human sensor input. Or it is behavior of human purposely.

**Acknowledgement:** This work is partially funded by a research grant from PRESTO, Japan Science and Technology Corporation (JST).



The concept of "Parasitic Humanoid"



The three components of the "SmartFinger"

## References

- [1] T. Nojima, M. Inami, Y. Kawabuchi, T. Maeda, K. Mabuchi and S. Tachi, An Interface for Touching the Interface, ACM SIGGRAPH 2000 Conference Abstracts and Applications, p.125, 2001
- [2] Mascaro S. and Asada H., 2001, "Photoplethysmograph Fingernail Sensors for Measuring Finger Forces Without Haptic Obstruction," IEEE Transactions on Robotics and Automation, Vol. 17, No. 5, pp. 698-708
- [3] T. Maeda, H. Ando, J. Watanabe, Y. Nonura and T. Miki, "A behavior modeling with wearable robotics -The study of Parasitic Humanoid (VI)", The 6th VRSJ Annual Conference, pp. 153-154 (Japanese)



# The Spatial Bi-directional Reflectance Distribution Function

David K. McAllister  
UNC, Dept. of Comp. Sci.  
davemc@cs.unc.edu

Anselmo A. Lastra  
UNC, Dept. of Comp. Sci.  
lastra@cs.unc.edu

Benjamin P. Cloward  
Vicious Cycle Software, Inc.  
bcloward@viciouscycleinc.com

Wolfgang Heidrich  
UBC, Dept. of Comp. Sci.  
heidrich@cs.ubc.ca

Combining texture mapping with bi-directional reflectance distribution functions (BRDFs) yields a representation of surface appearance with both spatial and angular detail. We call a texture map with a unique BRDF at each pixel a spatial bi-directional reflectance distribution function, or SBRDF. The SBRDF is a six-dimensional function representing the reflectance from each incident direction to each exitant direction at each surface point. Because of the high dimensionality of the SBRDF, previous appearance capture and representation work has focused on *either* spatial *or* angular detail, has relied on a small set of basis BRDFs, or has only treated spatial detail statistically [Dana 1999; Lensch 2001].

Although storing BRDF parameters in texture maps is not new, storing a unique, general BRDF at each pixel for hardware rendering has not been done since the BRDF model must have only a few coefficients but still have a simple enough equation to be evaluated in fixed-function hardware [Kautz 2000]. Our SBRDF representation stores the parameters of a Lafortune BRDF [Lafortune 1997] with a selected number of basis lobes at each pixel. This is only a little larger than an RGB texture map, but can represent a completely unique, highly general BRDF per pixel. The generality, compactness, and rendering efficiency of this surface representation make it suitable for use today in interactive applications and software renderers.

Our spatial gonioreflectometer, Figure 2, automatically acquires the SBRDF of real planar surfaces like those in Figure 1. Having four motors allows the device to capture anisotropic BRDFs. The motors move to different poses at a density of every five to ten degrees over the incident and exitant hemispheres. At each pose the device acquires a high dynamic range image. Data sets range from 150 to 8000 poses, and from one to 30 GB of image data.

To create an SBRDF from the measured data the system rectifies the measured images into the same image space then computes the BRDF parameters at each pixel based on the reflectance measurements at that point. The Lafortune BRDF is nonlinear, so we fit to it by partitioning the equation into a linear least squares problem within a line search. We also use Levenberg-Marquardt optimization, though it often requires manual iteration before converging and is slower than desired for automatic fitting of unique BRDFs at each pixel.

We render SBRDF surfaces in graphics hardware by first loading each SBRDF as texture maps: the diffuse albedo, the specular albedo, and the lobe direction and exponent. The entire scene renders in two passes, whereas current hardware BRDF rendering algorithms require passes for each unique BRDF [McCool 2001]. The first pass evaluates the Lafortune specular lobe at each pixel and stores it and the specular exponent in the frame buffer. The second pass reads the frame buffer as a texture, performs the exponentiation using a dependent texture read, multiplies the result by the specular albedo texture, adds the diffuse albedo, and finally multiplies by the irradiance from the light. These two passes are repeated for each light and each BRDF lobe. Details may be found in [McAllister 2002].



Figure 1. Real-time rendered SBRDF scenes. Top: Measured couch fabric, curtain, and plant leaf SBRDFs; synthetic wood and brass SBRDFs. 260,000 polygons, 3 lights, 57 frames / second on Nvidia GeForce 4. Bottom: Left: One texture map with  $2^{20}$  different BRDFs. Right: Fabric shows subtle shading changes due to variation in thread material and direction of anisotropy.

## References

- DANA, K. J., B. V. GINNEKEN, ET AL. 1999. Reflectance and texture of real-world surfaces. *ACM Transactions on Graphics* **18**(1): 1-34.
- LAFORTUNE, E. P. F., S.-C. FOO, ET AL. 1997. Non-Linear Approximation of Reflectance Functions. In *Proc. of SIGGRAPH 1997*, 117-126.
- LENSCH, H., J. KAUTZ, ET AL. 2001. Image-Based Reconstruction of Spatially Varying Materials. In *Rendering Techniques '01*, 103-114.
- KAUTZ, J. AND H.-P. SEIDEL. 2000. Towards Interactive Bump Mapping with Anisotropic Shift-Variant BRDFs. In *Eurographics / SIGGRAPH Workshop on Graphics Hardware*, 51-58.
- MCALLISTER, D. K. 2002. *A Generalized Surface Appearance Representation*. Ph.D. Dissertation. University of North Carolina. <http://www.cs.unc.edu/~davemc>.
- MCCOOL, M., J. ANG, ET AL. 2001. Homomorphic Factorization of BRDFs for High-Performance Rendering. In *Proc. of SIGGRAPH 2001*, 171-178.

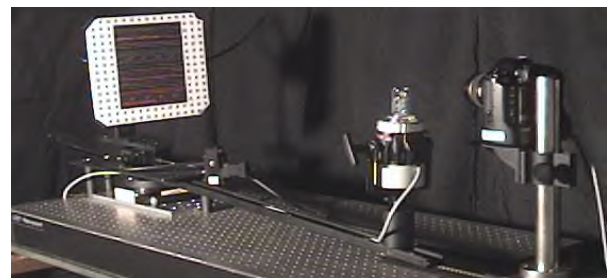


Figure 2. Acquisition device, with camera, moving light, pan-tilt-roll unit, and sample material with fiducial markers.

# The simulation of fluid-rigid body interaction

Tsunemi Takahashi

Toshiba Corp.

tsunemi.takahashi@toshiba.co.jp

Heihachi Ueki

Toshiba CAE Systems, Inc.

ueki@tci.toshiba.co.jp

Atsushi Kunimatsu

Toshiba Corp.

atsushi.kunimatsu@toshiba.co.jp

Hiroko Fujii

Toshiba Corp.

hiroko.fujii@toshiba.co.jp

## 1 Introduction

This sketch describes the modeling of the interaction between fluid and rigid bodies, how to simulate scenes in which fluid pressure acts on a rigid body, and conversely, in which rigid body motion drives a force back to fluid. We construct the interface between fluid simulation using the Cubic Interpolated Propagation (CIP) method and rigid body simulation using the impulse-based method. For fast simulation, we apply the CIP method to uniform structured meshes. For treating the interaction between rigid bodies and fluid efficiently, we use Volume of Solid (VOS) for rigid bodies, and for the collision among rigid bodies, we use Polygon-Polygon collision detection. For fast response to rigid body's collision, we use smaller time step for rigid body than for fluid.

## 2 Fast fluid simulation

For visually realistic motion of fluids, Navier-Stokes equations have to be solved. In the engineering area, the Arbitrary Lagrangian-Eulerian (ALE) method has been used for solving the fluid-rigid body interaction as moving boundary problems. But the ALE method uses unstructured meshes and requires high computational cost, and is unsuitable for applications in the computer graphics field. For a visually realistic scene and fast simulation, we choose the Volume of Fluid (VOF) [Hirt and Nichols 1981] method and the CIP [Yabe 1997] method for calculation of the advection term and apply these methods to uniform structured meshes. Using the VOF method, we can set boundary conditions easily. And using the CIP method, we can simulate fluid dynamics roughly but with sufficient visual reality.

## 3 Rigorous rigid body simulation

We employ two-level collision detection, Bounding Box Test and the Lin-Canny [Lin and Canny 1991] algorithm. Using Bounding Box Test, we check the possibility of the interpenetration among rigid bodies. If there are possibilities, we use the Lin-Canny algorithm to calculate the distance between them and the point where they collide. For collision response, we use the impulse-based [Mirtich 1996] method. Owing to use of polygon data for collisions among rigid bodies, we can check interpenetration of objects with high accuracy.

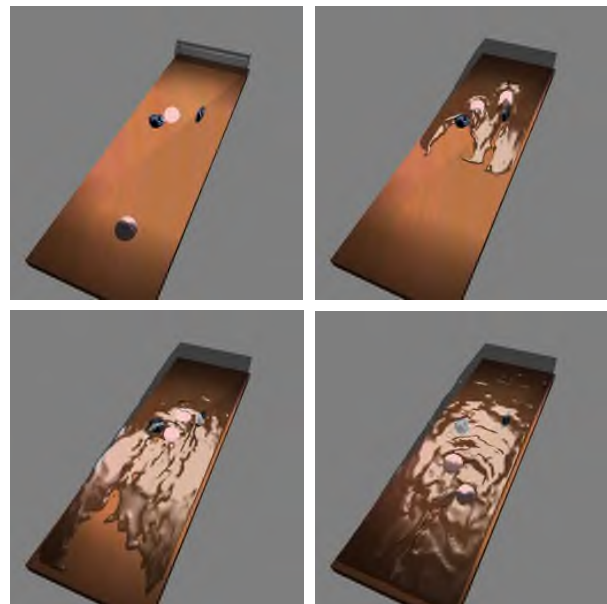
## 4 Fluid-rigid body interaction

Collision detection among rigid bodies is done with Polygon-Polygon collision detection, but the interaction between fluid and rigid bodies is done with VOS value. VOS value is the fraction of a solid's volume contained in a cell. We consider a cell with VOS value of more than 0.5 to be a solid cell, and consider any other cell to be a fluid cell. Boundary condition is set boundary between a solid and fluid cell. We make the gradient of pressure equal to zero along the boundary, and make the velocity of the fluid near the boundary fixed. This treatment may be too rough for use in the engineering field, but it provides sufficient visual reality. In order

to simulate the dynamics of rigid body, we calculate force and torque. For these, we consider only pressure of fluid.

## 5 Example and evaluation

The following figure shows one example of an animation generated by our system. There are two peaches rolling on the ground. Then the river-like flow appears and these peaches are floated away, colliding with rocks. A 120x36x18 mesh was used. We got scenes with consistent motion of fluid and rigid bodies, that is, very attractive and complex scenes that have not been achieved previously in computer graphics. The simulation time required in our system is acceptable for practical use. In view of several advantages mentioned above, our system is useful for practical application in the computer graphics field.



## References

- HIRT, C. W., AND NICHOLS, B. D. 1981. Volume of Fluid (VOF) Method for the Dynamics of Free Boundaries. *Journal of Computational Physics*, 39, 201-225.
- YABE, T. 1997. Unified Solver CIP for Solid, Liquid and Gas. *Computational Fluid Dynamics Review*.
- LIN, M. C., AND CANNY, J. 1991. A fast algorithm for incremental distance calculation. *International Conference on Robotics and Automation*, 1008-1014.
- MIRTICH, B. 1996. *Impulse-based Dynamic Simulation of Rigid Body Systems*. PhD thesis, University of California, Berkeley.

# The Time Geography Project - using computer graphics to visualize problems in social science

Kajsa Ellegård, Johan Torne, Henric Joanson, Anders Ynnerman, Matthew Cooper, Mark Ollila  
Norrköping Visualization and Interaction Studio  
University of Linköping  
Norrköping Campus, Sweden

In natural sciences the practice of using computers to visualize and analyze data was adopted early. In social and cultural studies, however, computer technology has not been as commonly used for visualization purposes. This project to visualize Time Geographical datasets, the Time Geography Project, is therefore somewhat groundbreaking.

Most social scientists claim that historical events and actions have great influence on the present situation and what will follow in the future. However, the necessary biographical approach is seldom used in a systematic way. The Time Geographic approach, developed in the 1950's and -60's by Professor Torsten Hägerstrand in Sweden, enables us to analyse people's continuous biographies in a day or a life perspective. In Sweden, for example, the church commenced the collection of data about each individual (birth, confirmation, marriage, migration etc) for tax purposes centuries ago. In an early application these data were used by timegeographers to explain land use patterns caused by emigration to the US.

Today, several studies to explain the time use in everyday life of men and women are being carried out, based on individual diaries. The results are presented as mean values for groups of individuals. Consequently, the richness of the diaries is not utilized for analyses and the complex patterns of activity sequences, project interruption etc are not taken into account. This is a pity since common sense makes us aware that the great difference between family life with a baby child and with a school child is not the total time the child claims from their parents but the way the claims are distributed over the day (and night).

This project examines the use of time geographical data to analyse people's everyday activities and movement in time and space in various environments. The research is based upon a dataset consisting, currently, of diaries written by 464 individuals in 179 households during two days and which consists of 29000 unique entries documenting 500 different types of activities. The activities are clustered by type: for example sleeping, saunas and brushing your teeth are all treated as 'care for oneself'.

The diaries are written by individuals as log books with, among other things, the start-time (indirectly revealing the end-time of the preceding activity), location and in the company of whom each activity is performed during the selected days. The common practice in Time Geographical studies is to draw individual paths based on the start-time and end-time of each activity. These paths were initially drawn on paper and then compared to find similarities and differences between selected small groups of individuals. This is a time consuming process since it is dependent on a person drawing each line either by hand or by computer and on that person's ability to find similarities between all the lines representing the group of individuals selected. If, however, we could use the computer to generate the paths within a virtual space where the user can navigate, this would make it possible for the user to concentrate their efforts on the correct selection of data and on the control of the navigation of the displayed data to zoom in and out on the lineset, rotate and focus on those parts which demonstrate the interesting features of

the data. The user could then make use of other options, for example, colour code different activities, locations, durations or 'company' in real time, to better emphasize those features for the audience. The use of a computer system would also allow the user to dynamically update the dataset to bring in other pertinent data or exclude that which is less significant. In our system we have imported the data into an SQL database which the visualization application can interrogate directly to extract data subsets on demand according to sets of selections made by the user. These data are then displayed within the application according to a range of parameters such as controlling the colour scheme to highlight specific activities, locations, companionship or the energy consumption associated with these activities. Other factors which can be determined by the user include simple graphical features such as the size and selection of the mapping of the data to the three axes.

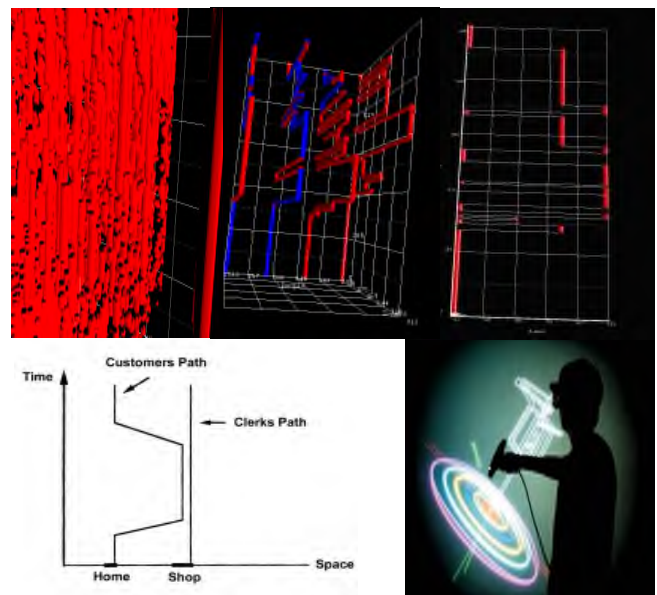


Figure: From top left moving clockwise. 1. Using visualization, the level-of-detail obtainable in viewing the population is enhanced. 2. The visualization features multicoloured tubes to represent each individuals activities during the course of a day. 3. This visualization shows an individual involved in an activity: baking a cake. The image illustrates time sharing over different tasks. Using the average would not show the interleaving of different activities which is clear from this visualization. Hence the statement "She spends 50% of her time in the kitchen" is not as meaningful as "She bakes the cake while watching TV". 4. Multimodal visualization capabilities: Using an alternative, polar coordinate representation on an immersadesk. The vertical-axis is time, radius (concentric circles) are different activities, with each individual assigned to a particular angle. 5. Traditionally time-geographers compare plotted individual paths to do analysis.



# Three Angry Men: Dramatizing Point-of-View using Augmented Reality

Blair MacIntyre<sup>1</sup>, Jay David Bolter<sup>2</sup>, Jeannie Vaughan<sup>2</sup>, Brendan Hannigan<sup>1</sup>, Emmanuel Moreno<sup>2</sup>,  
Markus Haas<sup>1</sup>, Maribeth Gandy<sup>3</sup>  
GVU Center, Georgia Institute of Technology

“Three Angry Men” is a novel augmented reality experience that explores the use of Augmented Reality (AR) as a dramatic medium. The user participates in an AR version of the famous twentieth-century play, “Twelve Angry Men,” [Rose 1983] which for practical reasons we have abbreviated into a scene involving 3 characters (thus, “Three Angry Men”). The participant finds herself immersed in a physical jury-room, where virtual characters (jurors in the drama, rendered as video-based characters overlaid at appropriate 3D locations around the physical table using a see-through head-worn display) debate the guilt of a young man on trial for murder (see Figure 1).

The goal of this experience is to dramatize the idea that different people perceive events differently, depending on their beliefs, backgrounds, prejudices, and so on. The participant is tracked as she moves around the table, and when she sits in one of the three chairs, she occupies the point of view of that juror, temporarily entering the character’s mind and hearing his inner thoughts. The six minute scene uses the same script for each point of view, but the experience is different: the inner thoughts of each character color the spoken text, and each character may appear strong or weak, neat or shabby, thoughtful or dithering, depending on the personality of the juror viewing them.

This experience is powerful, however, because it is more than just three different interpretations of the same script: the participant experiences the debate not as three linear scenes, but as a mixture of the three points of view. The participant directly controls the viewpoint by physically moving around the room; the scene pauses when she stands and continues from the appropriate *first-person* point of view when she sits in one of the three juror’s chairs. Thus, each participant’s experience is different, depending on how she moves around, and her perception of the debate is shaped by the combination of viewpoints that she experienced.

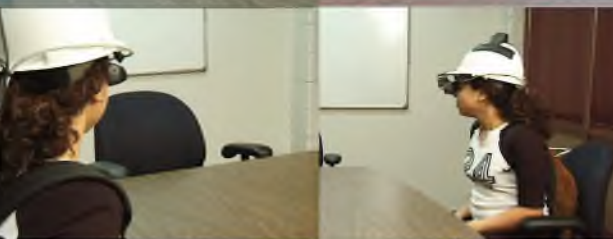
## Dramatic AR Experiences

This demonstration piece is an early product of our research program to develop AR technology for creating dramatic or narrative experiences with applications in education, the arts, and entertainment. Most researchers have explored AR as a straightforward “information-delivery” system, in which textual and other relevant information is added to the user’s field of view at the appropriate time. While we recognize the importance of such applications (we have built many ourselves over the past decade, such as [Feiner et al. 1993]), we believe that dramatic AR experiences also hold great potential. Our research group is working simultaneously to develop the new media theory and conventions needed to design such experiences and to build the authoring and runtime tools necessary to implement these designs.

Our method is to approach AR as a new medium, and examine it using media theory as one would any other medium (such as virtual reality, film or stage). In particular, we focus on remediating [Bolter and Grusin 1999] established media in an attempt to



(a) Jurors C and B, seen from each other’s view across the table.



(b) Juror A, seen from the other two juror’s point of view. The participant sees the juror from the viewpoint she physically occupies.

**Figure 1:** Three juror’s discuss the fate of a man accused of murder. At any point, the user may occupy the chair (and by extension, the mind and point of view) of one of the jurors. Each juror appears differently to the other two. For example, the older, conservative, caucasian juror (on the left above) sees the young caucasian juror differently than the young african-american juror sees him.

understand and develop media forms for AR. (Media forms can be thought of as sets of conventions and design elements that can be used by authors and developers to create meaningful information experiences for these users.) Media studies teaches us that remediation is a critical tool because we never design in a vacuum, even when designing for a new medium. A user’s expectations are (implicitly and explicitly) based on their experience with, and understanding of, all media forms; a lifetime of experiencing film, stage, tv and so on creates a starting point for their interpretation and understanding of any new experience. Understanding, and leveraging, these shared cultural expectations of the intended audience will allow us to create richer, more engaging and more understandable AR experiences.

## References

1. College of Computing, {blair,brendan,mhaas}@cc.gatech.edu.
2. School of Literature, Communication and Culture, jay.bolter@lcc.gatech.edu, jeannievaughan@earthlink.net, nmoreno@motive.design.net
3. Interactive Media Technology Center, maribeth.gandy@imtc.gatech.edu

- [1] BOLTER, J. D., AND GRUSIN, R. 1999. Remediation: Understanding New Media. Cambridge: MIT Press.
- [2] FEINER, S., MACINTYRE, B. AND SELIGMANN, D. 1993. Knowledge-Based Augmented Reality. *Communications of the ACM* 36(7): 53–62.
- [3] ROSE, R. 1983. Twelve Angry Men: A Play in Three Acts. Dramatic Publications Company.



# Towards Visualizing HCI for Immersive Environments: The Meta-Situational Tracker

Christopher Jaynes \*, Joan Mazur, Cindy Lio  
Metaverse Lab, Depts. Computer Science/Education  
University of Kentucky  
Lexington, KY 40506

## 1 Introduction

The Metaverse is a synthesized world that combines computer-generated elements with the real-world to either augment or supplant a user's view of reality. To fully participate in this meta-world, users must access it through a unique interface that is visually immersive and non-restrictive, interactive and collaborative. Our ongoing research program in immersive environments emphasizes techniques that make widespread deployment and use of the Metaverse feasible. To that end, we are addressing technical challenges related to self-configuring, self-monitoring displays. Other researchers have made significant progress in overcoming the obstacles related to building tilted displays. Although these technical advances have made low-cost immersive displays feasible for a larger community of users, there is little understanding of the remaining HCI obstacles that prevent widespread deployment of such systems. Our early work, as part of a five-year evaluation project, focuses on defining the methods that will be required to understand usability and utility of the technology. Additionally, we are developing tools that will assist researchers in characterizing the behavior of users within nonrestrictive immersive environments.

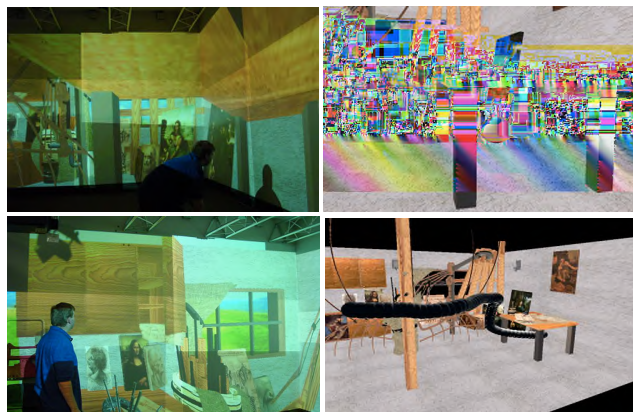
## 2 The CoRE Lab for Visually Immersed Behavior

Our initial studies are taking place in the Collaboratively Rendered Environments (CoRE) display but will be expanded to other environments over the next five years. The CoRE is a multi-projector, multi-PC display system that automatically calibrates the position of cameras and projectors, corrects for projector overlap and renders a seamless display on arbitrary surfaces. The display tracks users' motion within the environment and reacts to user head motion by modifying the rendered content appropriately. In contrast to fixed, high-end immersive displays, or head-mounted displays, the Metaverse project, as well as a number of similar research programs, promise to provide users with an immersive, nonrestrictive environment for relatively little cost and with little modification of infrastructure. The tools and evaluation methodologies we are developing will assist researchers in addressing important usability questions.

## 3 Evaluation Methodology and Tools

New methods for data collection are suggested by the unique spatial and temporal aspects of the Metaverse environment – the non-restrictiveness, for example. New situated evaluation techniques will be used to describe interactions, applications, and issues put forward by a large user community, and examine multiple points of view (researcher, designers, multiple backgrounds of users etc.). A transaction recorder, referred to as the *Metasituational Tracker*, is being developed to record physical and temporal user data that will supplement the use of more traditional usability methodologies such as interviews, observations, and survey data.

The Metasituational tracker provides a controllable visual display of a user's situation within the display environment either in real-time or in "playback" mode by loading a previously recorded session. Researchers can view the environment being visualized by a user, the user's position within the environment, statistical information about the visualization session (framerate, average user velocity, etc.). The tool is interactive and allows users to rotate the scene, as well as control the playback of previously recorded sessions. The transaction recorder outputs categories of data: 1) *Feature extraction* - what aspects of the environment are used and how? 2) *Feature evaluation* - What did the record show resulted from the action? 3) *Evidence Accumulation* - an aggregate of time spent doing certain actions and frequencies of certain actions, etc.



Example of the Metasituational tracker in use. (top left) User explores a 3D environment beneath a virtual table. (top right) Head position is recorded and presented as a sphere in the scene. (bottom row) User motion as a trajectory of viewpoints over time.

Preliminary research is needed to lay a sound foundation for suggesting areas for more systematic future research of immersive, non-restrictive environments. A thoughtful, empirical analyses of transaction records, using appropriate tools such as the metasituational tracker, will inform our understandings of human-computer interaction in new computer-generated environments where users' physical actions are symbiotically related to their exploration and understanding of complex information.

Jaynes, C., Webb, S., Steele, M. "A Scalable Framework for High-Resolution Immersive Displays", *International Journal of IETE, Special Issue on Visual Media Processing*, Arvind Nayak (eds.), 2002 (to appear).

Gay, G., Mazur, J. and Lentini, M. (1994). The use of hypermedia data to enhance design. *Computer Graphics: SIGGRAPH*, 28(19), 34-37.

Research funded in part by NSF grant No. 0101242

\* Email: Jaynes@metaverselab.org

# Tracking 3D Puzzle Pieces for Collaborative Learning Environments

Lori L. Scarlatos, Saira Qureshi  
Brooklyn College, CUNY  
Department of Computer and Information Science  
2900 Bedford Ave., Brooklyn, NY 11210  
lori@sci.brooklyn.cuny.edu

Shalva S. Landy  
CUNY Graduate Center  
Department of Computer Science  
365 5th Ave., New York, NY 10016  
landy@sci.brooklyn.cuny.edu

Children naturally learn about their world by manipulating objects within it. Playing with blocks and puzzles helps to develop their understanding of spatial relationships and other mathematical concepts. Using physical objects also allows them to work and learn in groups. Yet sometimes they need outside intervention from an adult or knowledgeable guide to help them learn more and stay engaged longer. Unfortunately, instructors often have too many students to give each one adequate attention. Our work focuses on developing computer-based “guides on the side” that can “watch” as children play with physical puzzles, and offer help or suggestions as needed. Our approach is to use the physical puzzle pieces as parts of a tangible interface. With our system, children are free to explore and collaborate without a computer, yet they can benefit from the computer’s instruction as they need it. We have successfully implemented and tested a 2D Tangram puzzle using this approach [Scarlatos 2002].

In this sketch we present our most recent work, which extends these ideas to the third dimension and applies them to a Soma Cube puzzle. This work is unique in that we are tracking multiple wireless 3D objects, in a small space, simultaneously. Here, we present two approaches to this problem. We have also developed a novel representation for the state of the 3D puzzle, which we describe. This representation enables our system to select appropriate hints, give encouragement as progress is made, and offer congratulations when the solution is found.

## 1. Seeing Puzzle Positions

Our first approach is to use computer vision to track the position and orientation of the puzzle pieces. Pairs of digital cameras capture stereo views in color. To help distinguish the pieces from one another, we painted them with distinct fluorescent colors (fig. 1). We selected these colors based on their distribution in YUV space. We then used reflective tape to mark the edges of the pieces. With the light source mounted near the cameras, these edges stand out in stark contrast to the rest of the environment.

We work with the images in two parts. First, we use the luminance (Y) values of the images to find the corners of the puzzle pieces. We achieve this with a Sobel gradient filter and Harris corner detection. Second, we use the chrominance (UV) values to determine which puzzle pieces the corners correspond to. We then correlate these corners in the stereo images to derive 3D coordinates. Finally, we translate, rotate, and scale the coordinates to achieve the best match with our internal representation of the puzzle piece, which yields its orientation [Trucco and Verri 1998].

## 2. Sensing Puzzle Positions

Our second approach is to use sensors within the puzzle pieces to detect their condition. We extended Anderson et al.’s approach for computational building blocks [2000] by using wireless communications to transmit information about the state of each puzzle piece. Each piece of the Soma cube contains a Basic Stamp, current sensors, and a radio frequency (RF) transceiver. A current sensor detects when a circuit is completed by two puzzle pieces touching (fig. 3). By placing distinct resistors at each possible circuit, we can tell which sides of which pieces are adjacent to each other. With additional resistors, we can detect relative orientation

as well. This adjacency information is transmitted to the computer via the RF transceiver. The setup of the leads on the outside of the pieces ensures that every adjacency will be sensed by the two pieces.

## 3. Representing Puzzle State

In order to respond appropriately to what the children are doing, we need a rotation- and translation-invariant representation of the state of the puzzle. We have two ways of representing this state. First, we represent each pair of touching puzzle pieces with a substring indicating: a) the labels of the pieces and their touching faces, b) a relative orientation angle, and c) their topological relationship (e.g. meets, intersects, contains). These resulting substrings are then sorted and concatenated to produce a unique string representing the precise state of the puzzle. Second, for puzzles that have multiple solutions, we represent clusters of touching pieces with a voxel model whose coordinate frame corresponds to the local coordinates of the piece with the lowest-valued index.

## 4. Acknowledgements

This paper is based upon work supported by the National Science Foundation under Grant No. 9984385, and by a PSC-CUNY grant.

## 5. References

- ANDERSON, D., FRANKEL, J.L., MARKS, J., AGARWALA, A., BEARDSLEY, P., HODGINS, J., LEIGH, D., RYALL, K., SULLIVAN, E., AND YEDIDIA, J.S. 2000. Tangible Interaction + Graphical Interpretation: A New Approach to 3D Modeling. In *Proceedings of ACM SIGGRAPH 2000*, ACM Press / ACM SIGGRAPH, New York, 393-402.
- SCARLATOS, L.L. 2002. TICLE: Using Multimedia Multimodal Guidance to Enhance Learning, *Information Sciences 140*, 85-103.
- TRUCCO, E. AND VERRI, A. 1998. *Introductory Techniques for 3-D Computer Vision*, Prentice Hall.



Figure 1. Single snapshot of Soma Cube puzzle pieces.

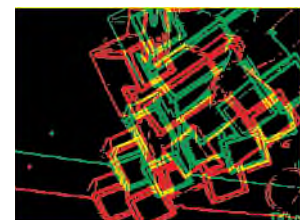


Figure 2. Edges detected in a stereo pair.

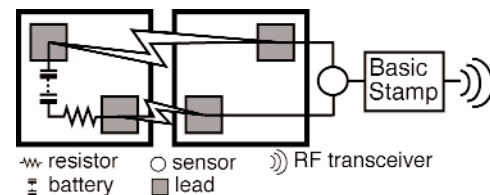


Figure 3. Touching surfaces complete the circuit.

# TWISTER: A Media Booth

Kenji Tanaka Junya Hayashi Yutaka Kunita Masahiko Inami Taro Maeda Susumu Tachi  
Graduate School of Information Science and Technology, The University of Tokyo  
e-mail: twister@star.t.u-tokyo.ac.jp  
http://www.star.t.u-tokyo.ac.jp/projects/mutel/

TWISTER (Telexistence Wide-angle Immersive STEReoscope) is an immersive full-color autostereoscopic display, designed for a face-to-face telecommunication system called “mutual telexistence”, where people in distant locations can communicate as if they were in the same virtual three dimensional space.

The basic idea of mutual telexistence is the projection of human beings into a virtual environment in real time [Tachi et al. 1996] [Kunita et al. 1999]. Each human user resides inside a booth, which displays a live full-color panoramic and autostereoscopic images to the user, and at the same time, captures images of the user from every angle. Autostereoscopic here means there are no obstacles to hide the user’s face, and therefore, the face of the observer can be clearly seen. With multiple booths, people at remote locations can meet together as if they were close at hand.

TWISTER has rotating display units surrounding the observer. One unit consists of two LED (Light-Emitting Diode) arrays and a parallax barrier. Each LED array consists of pairs of red, green, and blue LED’s, and displays time-varying patterns so that the observer can perceive an image. Due to the use of LED, TWISTER can be used as a display of high intensity full-color images for use in a bright room. The rotation of the display unit makes it a wide angle display. In fact, the angle of view as a normal display is 360 degrees.

The key device for autostereopsis is the rotating parallax barrier. One of the LED arrays is for the left eye, and the other is for the right eye. Because the parallax barrier obscures the opposite side LED emission, different images are shown to the left and right eyes. The angle of view as a stereoscopic display depends on the direction and the position of the observer. If the head of the observer is fixed, it exceeds 120 degrees. Thus, the observer can experience panoramic and autostereoscopic images without bothersome eyewear.

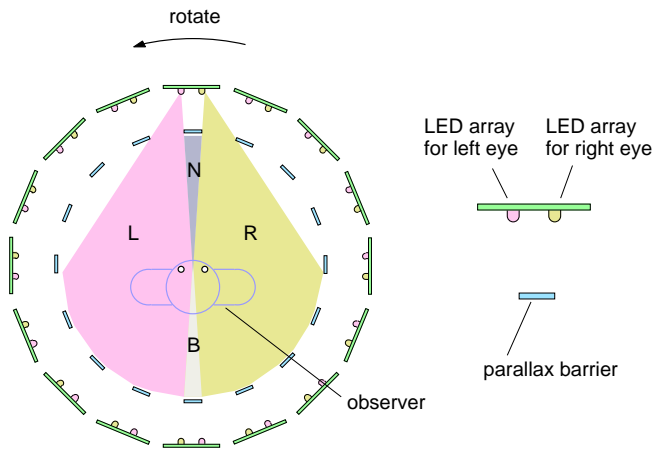


Figure 1: The principle of the rotating parallax barrier. In areas L and R, only images for the left and right eye can be seen, respectively. In area B, both images can be seen, while in area N no image can be seen.



Figure 2: A user stands inside a booth (left) and sees the figure of other person (right).

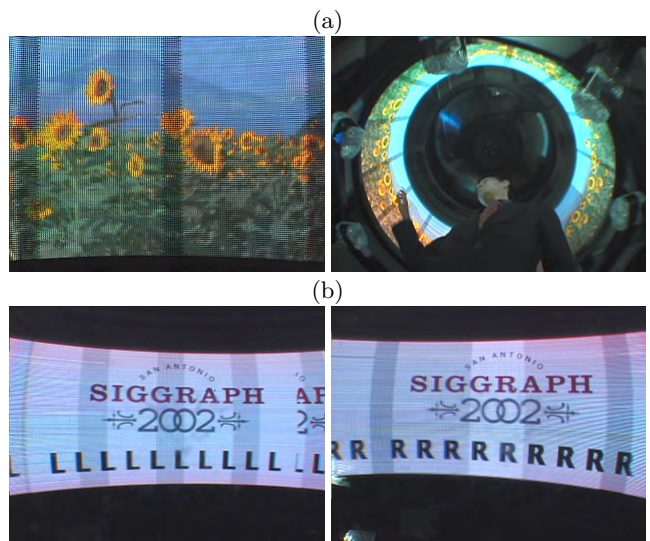


Figure 3: (a) The displayed full-color natural images. (b) TWISTER provides separate light sources to the left and right eyes.

*Acknowledgments* This research has been conducted as a part of the CREST Telexistence Communication Systems Project supported by JST (Japan Science and Technology Corporation)

## References

- KUNITA, Y., INAMI, M., MAEDA, T., AND TACHI, S. 1999. Prototype system of mutual tele-existence. In *ACM SIGGRAPH '99 Conference Abstracts and Applications*, 267.
- TACHI, S., MAEDA, T., YANAGIDA, Y., KOYANAGI, M., AND YOKOYAMA, H. 1996. A method of mutual tele-existence in a virtual environment. In *Proceedings of the ICAT'96*, 9-18.



# Uber Destruction in “The Time Machine”

Brad Parker  
Digital Domain

Digital Domain’s character animation team was asked to create a complex destruction sequence for Time Machine’s main villain... the Uber Morlock. This was to be no ordinary demise. Our villain had to age rapidly. His skin was to wrinkle, discolor, and disintegrate. His hair was to recede, his clothing to oxidize and fall apart. His muscles were to atrophy, eventually releasing his bones one by one. Finally his bones were to break apart and turn to dust.

We started with look development. Our art department created a series of illustrations portraying ten stages of decomposition. It was clear, from the approved designs, that seamless integration of our character, effects and compositing teams would be critical. R&D efforts began immediately on muscle deformation, an ‘Ebola’ type skin erosion, hair simulation, cloth simulation and bone decomposition.

Meanwhile, our on-set integration team was gathering survey data for camera tracking. Uber Morlock (played by Jeremy Irons) was Cyberscanned with and without costume for the modeling team. Camera match moving was done in Digital Domain’s proprietary, Academy Award winning tracking program.

Modeling of the outer skin surface was handled in Headus CySlice. Displacement maps were generated from the dense point cloud data and transferred, with custom tools, to a proprietary DD Subdivision Surface. Wardrobe, muscle and skeletal models were completed in Alias Wavefront’s Maya.

Setup and skinning of the character was handled in Maya. Muscle deformation was handled with custom deformer tools. Cloth was simulated using a series of proxy wrap deformers that linked to our RBD management tool. Hair was accomplished with a novel Maya/Houdini hybrid approach.

Each layer (skin, wardrobe, muscles, and bones) was available to the animation team through our Character GUI. DD’s character GUI handles auto versioning of character layers, shaders, and cameras. This approach saves time and effort allowing all teams to work in parallel while sharing the latest data.

Animation was completed in Maya. DD’s animators were required to create a believable performance throughout the sequence. A combination of facial blend shapes and deformers were used to express pain and distress. Final animations drove the hair, cloth and rigid body simulations.

A custom Rigid Body Dynamics management tool was developed for creating large simulations in Maya. This tool allows the user to break a large RBD simulation into manageable layers. Each layer could be independently edited. This tool also allowed for intermediate simulation objects, cutting down on simulation time while not sacrificing quality.

Digital Domain’s data management pipeline was used extensively in Time Machine. Geometry, UV coordinates, cameras and motion were all transferred seamlessly between Maya and Side Effects Software’s Houdini.

A procedural approach was used for surface decay in Houdini. Data sets were seamlessly transferred from Maya and organized for shading in Houdini. The Uber Morlock skin tearing and eroding effect was based on a system of reading animated color maps into a surface shader and using the rgb information as inputs that discolor and punch holes into the skin geometry. Each of the three colors would serve a different purpose in the shader for the purpose of describing where and when tears occurred. Hole generation for the bones needed to have a thickness (unlike the skin). A geometric solution was done using Houdini Sops. The principle was the same; unfold into "tear" space and apply coordinates, refold back into original form, and use coordinates to "erode". The difference was that a new model would be created for each frame. Cracks were generated by animating L-Systems and using Copy Stamps to scatter and rotate them around. Finally, each bone was turned to dust using particle systems birthed from each tear map.

Rendering was done in Side Effects Software’s Vmantra with the exception of hair, which was completed in Pixar’s Photorealistic Renderman. High Dynamic Range photos taken on set were used to generate lighting information and individual lights were added to that to enhance certain regions of the character. Over forty layers were rendered in RGBEA extended color space and composited in Digital Domain’s Academy Award winning compositor, Nuke.

Many techniques were used to create the Uber Morlock destruction sequence including but not limited to:

- Proprietary subdivision surface modeling and displacement generation.
- Rigid Body Dynamics with custom management tools.
- Muscle simulation.
- Hair simulation and rendering.
- Cloth simulation.
- Houdini particle systems for bone to dust generation.
- Houdini L-Systems.
- High Dynamic Range rendering.
- Extensive inter-package data management.
- Lifelike character animation.
- Photorealistic rendering of a humanoid character that had to match perfectly with live action material.

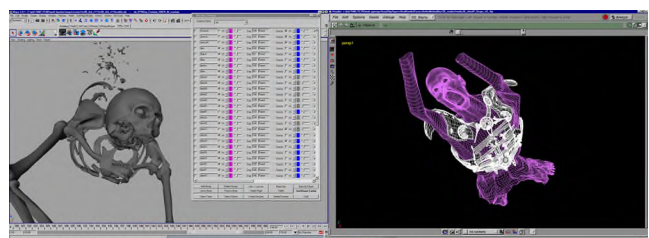


Figure 1. (a) Custom RBD management tool in Maya  
(b) Wireframe view of Uber Morlock in Houdini



# User Customizable Real-Time Fur

John Isidoro  
Boston University and ATI Research  
Jisidoro@ati.com

Jason L. Mitchell  
ATI Research  
JasonM@ati.com

## Introduction

Recent advances in real-time fur rendering have enabled the development of more realistic furry characters. In this sketch, we outline a number of advances to the shell and fin based fur rendering technique by Lengyel et al [2001] using the pixel and vertex shader capabilities of modern 3D hardware.

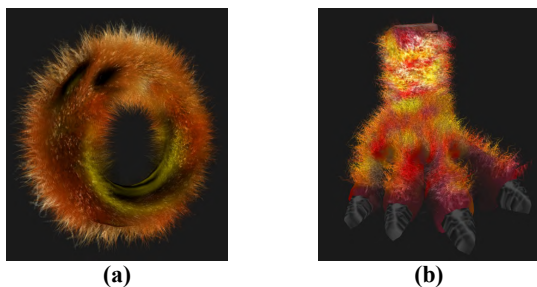


Figure 1 – Furry torus and animal claw

## Shell and Fin extrusion using Vertex Shaders

Vertex shaders allow for more flexibility in fur rendering by performing the shell and fin extrusion in hardware. For shells, each shell vertex is moved in the direction of the normal a fraction of the fur height. For each of the fins, two of the four vertices are extruded in the same manner. The fin's  $v$  texture coordinate is used to choose which vertices get extruded. One of the nice things about extrusion in a vertex shader is that it allows for the fur length to be set using a vertex shader constant. A possible application of this is to have characters in a video game have shrinking or growing hair during the course of a game.

## Per Pixel Anisotropic Lighting on Fins and Shells

One interesting property of stochastically generated fur is the distribution of hair directions throughout the texture. Due to the changing directions of curly hairs, it is no longer sufficient to derive hair direction for lighting from only the albedo texture. Instead, the hair direction is encoded per pixel along side the opacity value when the 3D shell and 2D fin fur textures are generated. When rendering the fur, the per-pixel direction can be used as a direction of anisotropy for Heidrich and Seidel's [1999] strand anisotropy technique. This allows individual hairs to capture diffuse and specular lighting. This helps to break up the homogeneity of the fur, producing a more realistic rendering.

## Hair Color Sourcing (Shells)

Another technique we use to create more realistic fur is to apply an albedo texture that produces color variations for different regions of the fur. The alpha channel of this texture can also be used as an opacity value to produce bald spots on regions of the model. However, when the fur texture contains very wavy or slanted fur, each hair will change color over its length due to the fact the texture coordinate used to fetch from the albedo texture is directly tied to the hair's position. A better approach is to fetch the same texel from the albedo map over the length of the hair. This way, when there are sharp edges in the albedo map, the hairs from both sides of the edge will intermingle, giving a more realistic appearance. For example, the hairy foot in Figure 1b does not have fur on the claws. The hairs originating near the claws are able to bend over the claws while remaining visible and colored correctly.

## Color Sourcing for Shells

To implement color sourcing for shells, another 3D texture is generated in parallel with the direction/opacity texture. This texture contains a  $(u,v)$  offset for each pixel with the hair which represents how far the  $(u,v)$  position of the hair is from the  $u,v$  position of the albedo of the hair. In the pixel shader used for rendering the shells, the albedo texture lookup is perturbed by this  $(u,v)$  so that each hair fetches from only one position in the albedo map. This way, each hair's albedo is the same over the length of the hair.

## Color Sourcing for Fins

Color sourcing for the fin texture is performed in a similar manner, but requires additional per vertex data in order to perform the correct math. Because the fin fur texture is 2D, only a  $u$  offset is required per pixel. However, the fin fur texture coordinates are different from the albedo map texture coordinates, and an albedo map  $(u,v)$  needs to be computed from the fur texture's  $u$  offset. To do this, the albedo map  $du, dv$  for each fin is encoded per vertex, and interpolated over the polygon within the pixel shader. In the pixel shader, the fur texture's  $u$  offset is multiplied by the albedo map overall fin  $du, dv$  in order to get the hair's  $(u,v)$  offset for the albedo map.

## Controlling Hair Density Using the Albedo Map

Another parameter that can be specified over the surface of the object using an albedo map is hair density. In order to do this, an additional texture channel is generated in parallel with the other fur texture channels. In this channel is a "thinning" value from 0.0 to 1.0 which specifies the percentage of the fur texture left to be generated. The first hair generated gets a value of 1.0, and the last hair generated gets a value of 0.0. In the pixel shader, a pixel is only written if its thinning value is less than or equal to the albedo map's density value (usually stored in the alpha channel of the albedo map). This density value causes a percentage of the hairs to not be drawn, thus thinning the fur.

## Hole Filling for Mipmap Generation

Simply down-sampling the color-encoded direction and offset textures causes incorrect results due to the unfilled pixels in the texture. In order to get around this, each unfilled pixel in the texture is replaced by a weighted average of the filled pixels within a small kernel radius. This results in a texture which varies smoothly, and can be down-sampled without producing artifacts in the lighting or albedo map offsets.

## Future Enhancements

Since both the fins and shells are extruded along the direction of the vertex normal in the vertex shader, it is possible to animate the fur by perturbing the normals. This perturbation could be performed using a vertex tweening sequence or procedurally inside a vertex shader.

## References

- HEIDRICH, W. AND SEIDEL, H.-P. Realistic, Hardware-accelerated Shading and Lighting. In *Proceedings of SIGGRAPH 1999*, ACM Press / ACM SIGGRAPH, New York. 171-178.
- LENGYEL, J., PRAUN, E., FINKELSTEIN, A. AND HOPPE, H. Real-Time Fur over Arbitrary Surfaces. In *ACM 2001 Symposium on Interactive 3D Graphics*, 2001.

# Using Cellular Phones to Interact with Virtual Environments

Bent Dalgaard Larsen    Jakob Andreas Bærentzen    Niels Jørgen Christensen  
Technical University of Denmark\*

Distributed VR promises to change the way we work and collaborate [Singhal and Zyda 1999]. In this sketch we will extend the accessibility of the virtual world originally developed in [Larsen and Eriksen 1998] by introducing the use of the modern cellular phone as a platform for primitive interfaces to VR applications. We believe that our use of a cellular phone has led to the first completely pocketable platform for VR user interfaces.

While a cellular phone will offer only primitive access to the virtual environment, we expect that being able to access the VE at all times will spawn new and enhance existing applications of VR. For instance, existing online multi-user games may provide limited access through a cellular phone enabling people who either in transit or just not near a PC to participate through a more limited interface. An important question is “how limited?”. We believe that a “cellular interaction mode” need not be very restrictive as long as careful use is made of the technology available.

Unfortunately, existing cellular phones have no hardware support for 3D graphics and the screen is much smaller (Typically  $176 \times 208$  or  $208 \times 320$  pixels). Furthermore, the only input is from a limited number of push buttons. But a cellular phone has a high-speed TCP/IP connection to any server on the Internet using the current GPRS (General Packet Radio Service) protocol (40.2 kb/s) or the upcoming and much faster UMTS (Universal Mobile Telecommunications System) protocol.

Currently, the producers of cellular phones have chosen no common operating system standards. However, many of these mobile devices support a subset of Java (J2ME—Java 2 Micro Edition). This contains no floating-point math and the only supported network protocol is HTTP which implies that the only way to obtain duplex data streams on the cellular phone is if it continuously polls the server.

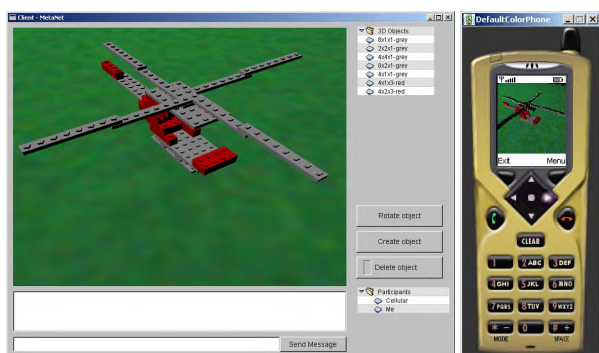


Figure 1: Interface on the workstation and the cellular device.

We have chosen to construct a LEGO™ building application and our task was to enable people to build LEGO™ structures in collaboration using either a traditional workstation or a cellular phone.

The system is based on a client-server model. The server contains the 3D models and maintains connections to clients. In our case, the clients may be either workstations or cellular devices. In the following discussion, we focus on the case where the client is a cellular device.

\*e-mail: (bd1|jab|njc)@imm.dtu.dk

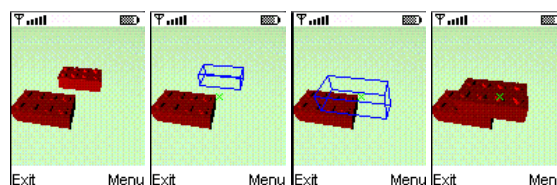


Figure 2: Local modifications on the cellular phone.

The lack of support for a 3D graphics API and floating point arithmetic makes it very important to divide work between the server and the cellular device in an intelligent way. In our system, we have chosen to perform high quality rendering on the server which uploads images to the device on demand. Since LEGO™ bricks manipulated interactively cannot be rendered locally in real-time, they are rendered in wire-frame. In the absence of floating point numbers, we have implemented a simple facility for fixed point arithmetic.

As the user moves in the environment new images on the server are generated and transferred to the cellular phone. When the user selects a LEGO™ brick by pointing at it, the server sends a subimage (the changed part of the image) to the user without the selected brick, which is then drawn in real-time wire-frame while the user alters its location and orientation. As the user deselects the brick, another subimage is generated on the server and the high quality image on the mobile phone is then update again.

The user interface is implemented as follows:

- A locator is shown on the display of the phone, which is controlled by the arrow keys.
- When the selection key is pressed, the brick under the locator is selected using a ray-box intersection algorithm.
- The brick is moved using the arrow keys. The arrow keys are used to move the brick in the ground-plane. The height value is found by moving the brick the smallest possible distance from its previous height value, given the constraint that the brick does not intersect other bricks.

Users located in the same VE and connected using a workstation have the same possibilities of manipulating LEGO™ bricks. Modifications made either through the cellular phone or the workstation are instantly reflected on all devices connected to the virtual environment.

Although this is still work in progress, our conclusion is that mixing locally rendered images on the cellular phone with server rendered images is an appropriate solution for applications such as ours.

This work was supported in part by the STVF project DMM and the Nordunit2 project NETGL.

## References

- LARSEN, B. D., AND ERIKSEN, M. A. 1998. *Multi-User Interactive Virtual Environments*. Master's thesis, Technical University of Denmark.
- SINGHAL, S., AND ZYDA, M. 1999. *Networked Virtual Environments*. Addison Wesley.

## Using the Virtual Terrain Project to Plan Real Cities: Alternative Futures for Hangzhou, China

Dr. Michael Flaxman,  
Lecturer in Landscape Planning  
Harvard Graduate School of Design  
mflaxman@gsd.harvard.edu

### Introduction

Traditional architectural software tools are well suited to individual building design, but don't scale up well to large landscapes and cities. In city and regional planning applications, it is important to visualize city blocks and neighborhoods, a task which requires integration of data from geographic information systems and 3D surface models from CAD. However, the development of semantically and visually rich real-time representations of cities and large landscapes has proven difficult. Two major problems are the acquisition of appropriate data, and the volume of data once acquired. A typical city or suburban landscape may have thousands of streets, tens of thousands of buildings, and hundreds of thousands of trees and bushes. Three dimensional data specifically describing all of these objects is not usually available, and visual simulations must be based on abstract classifications originally developed for entirely different purposes, such as tax assessor's databases or generalized vegetation maps.

Fortunately, several powerful data management techniques are available, including procedural modeling based on GIS attributes and perceptually-based level of detail (LOD) management. While both approaches have been known for some time, advances in technical implementation and hardware have only recently made it possible to generate real-time visualizations of large areas using standard personal computers. The key is to automatically generate the appropriate level of detail from large geographic data sets using representational techniques specific to the class of object and its perceptual characteristics.

The Virtual Terrain Project<sup>1</sup> (VTP) is an open-source software project whose goal is to foster the creation of tools for easily constructing any part of the real world in interactive, 3D digital form (Discoe, 2002). VTP software is cross platform and requires no specific hardware other than a good OpenGL graphics card. It can render terrain with continuous level of detail simplification, and provides switchable LOD representation for both surface vegetation and buildings. Also, VTP can import data from an enormous variety of CAD and GIS data formats, even converting between multiple spatial coordinate systems.

### Demonstration

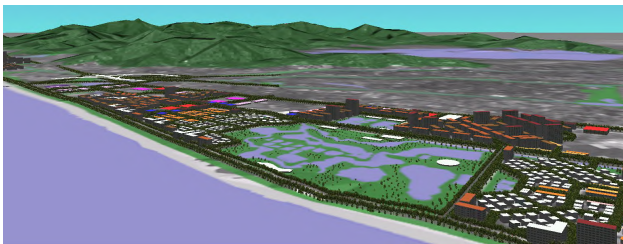


Figure 1: Screen capture from real-time visualization showing the location of a proposed wetland park relative to the West Lake (upper right).

This sketch will walk through the steps necessary to generate a virtual world using VTP. The software and sample data sets will be made available to participants in the Creative Applications section. Both binaries and source code are also downloadable from the web.

### Case Study

As an example, I will show a visualization of Hangzhou, China created by students at Harvard's Graduate School of Design<sup>2</sup>. The city of Hangzhou is located approximately two hours west of Shanghai. It is famous for the mountain-backed West Lake, one of the classic landscapes of China. It is also a major urban center (population ~6 million). The city is growing rapidly, and the lakeshore is under enormous development pressure. In the fall of 2001, the city of Hangzhou commissioned a study of alternative futures for the city which might refocus growth away from the lake and along the Xiantong River to the South.



Figure 2: Custom CAD models (Opera house and bridge) and buildings procedurally generated from GIS attributes. Roof colors represent land use type.

A conventional approach would have been to design and present a project of this magnitude using maps delineating relatively large homogenous zones separated by dominant land use. However, the studio wanted to use urban design concepts based on mixed use, and both the vertical dimension and the visual relationships between elements were judged to be critical. For example, the views from the West Lake are among the most treasured in China. In design terms, the challenge was to protect these views and to propose new plans for a city center which must span a very wide and dangerous river. The Xiantong is nearly a mile wide, and is famous for its tidal bore. Many design strategies involved using large civic structures as visual anchors.

The combination of the tools and strategies described above were successful in allowing students to generate four practical working models of the city, all within a one-month period.

1. Discoe, B. The Virtual Terrain Project website.  
<http://www.vterrain.org>

2. Steinitz, C. et al. Alternative Futures for Hangzhou, China. Harvard Graduate School of Design report. 2002.



# Video Textures Using the Auto-Regressive Process

Neill Campbell \*, Colin Dalton, David Gibson and Barry Thomas  
University of Bristol, UK

## 1 Introduction

Recently, there have been attempts at creating ‘video textures’, that is, synthesising new video clips based on existing ones. Schodl et al. showed new video clips by carefully choosing sub-loops of an original video sequence that could be replayed.

We present a different approach to recreating (potentially infinitely long) new sequences. We transform each frame of the video into an eigenspace using Principal Components Analysis (PCA) so that the original sequence can be viewed simply as a signature through this low-dimensional space (see Gibson et al. for another example of using PCA to assist animators). A new sequence can be generated by moving through this space and creating ‘similar’ signatures. This similarity is derived using the auto-regressive process (ARP) as discussed by Blake and Isard. A 2nd-order process is used and provides a statistical framework for assessing the quality of the new sequence. The new signature has the 2nd-order properties of the original and is much more than simply a random-walk through the space. New sequences created with our approach can contain images never present in the original sequence and are very convincing.

## 2 Results

Our first example is that of a campfire sequence. A plot of the original signature is shown below in red, with the synthesised signature superimposed in blue. The sequence is reconstructed using eight eigenvectors (using fewer may lead to blurring) and well characterises the flickering of the fire.

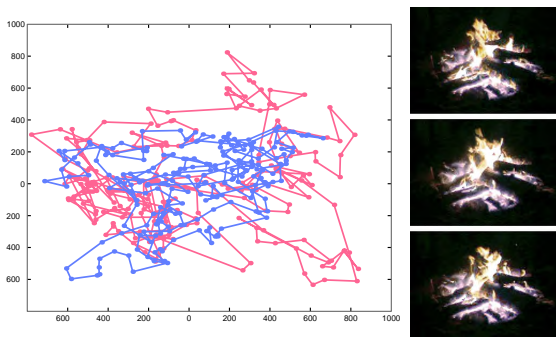


Figure 1: (Left) The original campfire sequence (red) vs the newly synthesised sequence (blue). For clarity, only the 2 dominant eigenvectors are shown. (Right) three of the resulting frames from the new sequence.

The ARP assumes that the signature is a Gaussian process, but sometimes image sequences have highly non-gaussian characteristics. An example of this are candle flames, where, due to their bending, correlations between images are not well explained using a linear analysis technique such as PCA. To overcome this we use the Combined Appearance Model of Cootes and Taylor to decompose the appearance into a shape description (a polymesh) and

a texture map (warped into a neutral pose). The combined PCA space is treated by us in the same way as the image eigenspace used for the campfire sequence, but exhibits the Gaussian nature required by the ARP. Shown below is a synthesised sequence of a man laughing created using such an appearance model. Once again the sequence captures the feeling of the original sequence well with the man smiling, laughing, shaking his head and rocking back and forth. The minor distortions present are due to the image warping technique used and we will overcome these. A second non-gaussian example fits a spline through the data space a synthesises a horse running on a treadmill using the distance along the spline, and an offset from the spline.



Figure 2: Four frames of laughing man. All of the images shown here are new and were not present in the original sequence.

## 3 Conclusion

We have presented a novel approach to creating video textures of arbitrary length. Such technology has a wide applicability, including the automatic generation of computer backgrounds and crowd scenes. This image-based rendering technique has been demonstrated for video sequences and also combined appearance models. The signatures on which the resulting images are based have the same properties as the original signatures and the sequences capture the look-and-feel of the original ones.

\*e-mail: Neill.Campbell@bristol.ac.uk



# Virtual Human Interface: Towards building an intelligent animated agent

Bernadette Kiss  
VerAnim, Budapest  
[BernadetteKiss@digitalElite.net](mailto:BernadetteKiss@digitalElite.net)

Gábor Szijártó  
VerAnim, Budapest  
[Szjarto.Gabor@freemail.hu](mailto:Szjarto.Gabor@freemail.hu)

Barnabás Takács\*  
Digital Elite, Los Angeles  
[BTakacs@digitalElite.net](mailto:BTakacs@digitalElite.net)

## Abstract

We describe our ongoing research on creating a *Virtual Human Interface* that employs photo-realistic virtual people and animated characters to provide digital media users with information, learning services and entertainment in a highly personalized manner. Our system was designed to be able to create emotional engagement between the virtual character and the user, thus increasing the efficiency of learning and/or absorbing any information broadcasted through this device. We developed innovative technologies for (i) photo-real facial modeling & animation, (ii) context dependent motion libraries with on-line retargeting, (iii) artificial emotions to modulate the characters' behavior and (iv) artificial vision to make the virtual human "aware" of its surroundings. The second key aspect of our solution is a simple to use high level content authoring process, comprising of video-based MPEG4 facial tracking and an innovative interface called the "Disc Controller", which allows users to create new actors, make them move and even direct them to achieve a final rendered output within minutes.

**Keywords:** digital actor, virtual human, synthespian, facial modeling, real-time animation, intelligent & autonomous agents, AI, Artificial Emotion, MPEG4, motion retargeting.

## 1 Overview

To meet the challenge of creating a believable animated human agent we argue that the key "catalyst" in the process of transforming information to knowledge lies in the system's ability to modulate emotions in the user. Thus, the content being broadcasted ("WHAT") should be independently and intelligently modulated by the emotion channels ("HOW") and delivered in a visually rich, but very focused - instead of a distracting - environment ("WHERE"). This triad of "what", "how" and "where" forms the platform for a new kind of communication paradigm that increases efficiency in absorbing information of any kind, and can be used for personalized and interactive education, communication and marketing.

## 2 What, How, and Where

Our system is an innovative content authoring and production tool that can be used in a stand-alone mode, operated over the Internet or create input for animation packages such as Maya and 3D Studio Max. The "What" channel refers to the script, educational and/or marketing content the producer wishes to deliver to a target audience. The creative process of "How" comprises of two major components. The first one focuses on designing a deterministic behavior by allowing users to create a photo-real, high resolution face (*Face Designer*) and animating the resulting head via facial tracking data via *MPEG4* streams as well as the body from motion capture libraries. The second component allows the virtual actor to deliver this animation either deterministically or effected by interactions from the user or the internal state of the actor as modulated by its 3D environment. Finally, the "where" channel creates a 3D world surrounding the actor in a real-time environment and mixes the objects of this world with that of the

user thereby enhancing the perception of reality. This enhancement is achieved by the combined effect of (i) controlling the motion of the main camera via *facial tracking* and/or external devices, such as virtual reality (VR) glasses. (ii) Allowing the virtual human to *see, detect and recognize* one or multiple people in front of the display and seeing them via a monitor much the same way the user sees the digital actor.

## 3 Virtual Human Interface System Highlights

**Facial modeling pipeline:** Turns a large set of 3D facial reference scans into animatable, high-fidelity models capable of expressing refined shades of emotions and subtle details to modulate content.

**Face Designer:** A database storing an extensive set of up to 60 different expressions for a large number of individuals. High resolution (30-50,000 polygon) models of various race, gender age and ethnicity are used to mix new heads in a matter of seconds.

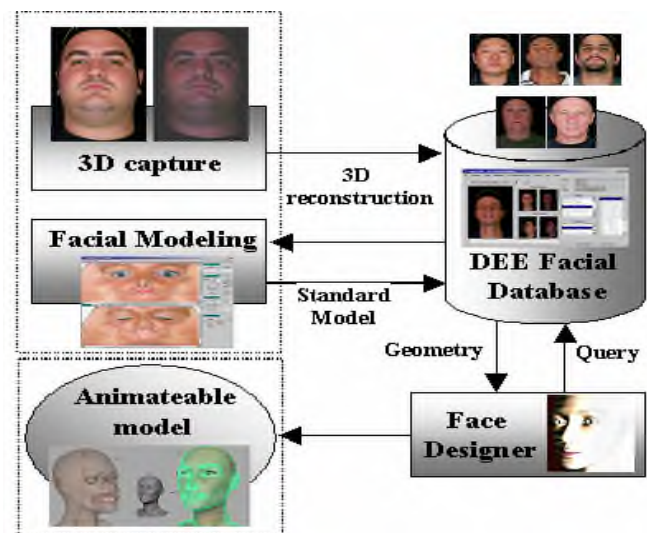
**Speech & Facial Animation:** The speech animation module accepts audio streams for basic LipSynching, MPEG4 compatible facial motion data and/or facial MOCAP streams each automatically retargeted seamlessly to any face.

**Motion Retargeting & Libraries:** Motion capture input as well as motion libraries support the fast creation of animated events. A simple interface (*Disc Controller*) is used to retarget motion in real-time and add emotional context to the original sequence.

**External sensory input:** The virtual human is aware of its environment and possesses senses of seeing, hearing, and touching that modify its behavior. It uses *Face Recognition* to detect, recognize and track multiple users in front of the screen and communicate with them attentively.

**Artificial Intelligence and Emotions:** Internal models of behavior as well as input received from its environment effect the digital actor's emotion and mood and used to bias the delivery of the original content.

**Multiple Animation Platforms:** Models and animated sequences can be view using the standalone real-time platform, over the Internet, or exported to standard formats to be used with animation packages such as Maya or 3D Studio Max.



# Virtual Reality Interfaces Using Tweek

Patrick Hartling\*, Allen Bierbaum†, Carolina Cruz-Neira‡  
Virtual Reality Applications Center (VRAC), Iowa State University

## 1 Introduction

Developers of virtual environments often face a difficult problem: users must have some way to interact with the virtual world. The application developers must determine how to map available inputs (buttons, gestures, etc.) to actions within the virtual environment (VE). As a result, user interfaces may be limited by the input hardware available with a given virtual reality (VR) system.

To address such limitations, we have developed Tweek, a middleware tool that presents users with an extensible Java graphical user interface (GUI) capable of communicating with VR applications. Tweek can run on desktop computers, on palmtop computers in projection-based virtual reality systems, or in a three-dimensional virtual space. We use Tweek to address some limitations of input and interaction within virtual environments (VEs) by designing GUIs that utilize familiar two-dimensional (2D) components. Interaction techniques that require fine-grained control or high-precision input are often better suited to 2D interfaces than the use of gestures or buttons on a wand [Hill 2000].

The use of Java-based GUIs for interaction within VEs has been investigated in previous works at the VRAC. The first such work investigated the usability of a palmtop system with a Java GUI to interact with and manipulate a virtual space [Hill 2000]. Tweek extends this research by generalizing the interaction capabilities and the dynamic GUI component loading.

Tweek is a collection of multiple technologies: C++, Java, JavaBeans, and CORBA. Combined, these allow a Java GUI composed of plug-ins to communicate with a C++ application. Our implementation aims to simplify the inter-language communication so that programmers can make use of Tweek in their VR applications without knowing all the details of the individual technologies.

At the heart of the inter-language communication is CORBA, the Common Object Request Broker Architecture [OMG 2001]. It provides a cross-platform, language-independent method for distributed objects to communicate. In Tweek, it manages all communication between the Java GUI and the C++ VR application. Because CORBA is language-independent, there exists the potential for use of other languages besides Java and C++.

The Java GUI itself is a generalized framework that loads JavaBeans [JavaSoft 1997] dynamically using XML-based descriptions. The JavaBeans may encapsulate any functionality, but those that extend the GUI are crucial to the use of Tweek in a VE. Such graphical JavaBeans are written by the VR application developers and are customized for use with a given application. Dynamic extension of the GUI allows the VR application to add components to the interface while the user is active within the virtual space. Through dynamically extensible input, we can explore new possibilities for interactivity in VEs.

## 2 Implementation

Tweek utilizes an extensible object-oriented design based upon the Observer design pattern [Gamma et al. 1995]. The subject is the

state of the VR application, and the observer is the GUI. The state information is specific to each application.

An example use of Tweek on the desktop is shown in Figure 1a. Here, a top-down navigation tool is available for point-and-click navigation within a three-dimensional (3D) space. The observed state is the current user position in the VE. Bookmarks can be selected for direct navigation to known locations, and the 2D map can be clicked to navigate to the selected location. The same navigation tool can be used in an immersive environment using a palmtop computer, shown in Figure 1b.

## 3 Conclusion and Future Work

At VRAC, we use Tweek on palmtop computers to control VR applications in our projection-based VR systems. To map Tweek into a 3D environment, we plan to experiment with existing work such as 3Dwm [Elmqvist 2001]. We hope to offer an immersive interface with familiar controls without re-inventing GUI technology.

## References

- ELMQVIST, N. 2001. *3Dwm: Three-Dimensional User Interfaces Using Fast Constructive Solid Geometry*. Master's thesis, Chalmers University of Technology, Göteborg.
- GAMMA, E., HELM, R., JOHNSON, R., AND VLISSIDES, J. 1995. *Design Patterns: Elements of Reusable Object-Oriented Software*. Addison-Wesley Professional Computing Series. Addison-Wesley Publishing Company, New York, NY.
- HILL, L. C. 2000. *Usability of 2D Palmtop Interaction Device in Immersive Virtual Environments*. Master's thesis, Iowa State University, Ames, IA.
- JAVASOFT. 1997. *JavaBeans 1.01 Specification*. Sun Microsystems, Mountain View, CA.
- OMG. 2001. *The Common Object Request Broker: Architecture and Specification*, 2.6 ed. Object Management Group.

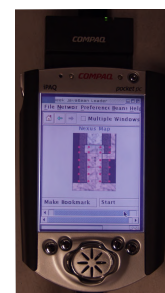
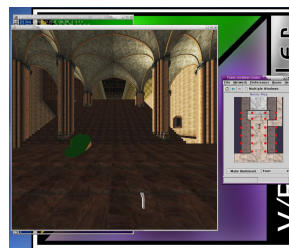


Figure 1: Tweek navigation on desktop (a) and PDA (b)

\*email: patrick@vrac.iastate.edu

†email: allenb@vrac.iastate.edu

‡email: cruz@iastate.edu

# Virtual View Generation by Linear Processing of Two Differently Focused Images

Akira Kubota\* and Kiyoharu Aizawa

University of Tokyo, Department of Electrical Engineering

## 1 Introduction

In this sketch, we present a novel approach to image-based rendering (IBR) techniques for generating a virtual view image from different positions with arbitrary focus using two differently focused images captured from a fixed position. In the conventional approach [Potmesil and Chakravarty 1981], focus blur have been produced by applying a realistic camera model to given 3D objects. We propose here a much simpler and more effective method only using space-invariant filters to render both parallax and focusing effects on objects in a scene, according to the virtual camera's position and focus depth respectively. The proposed method does not need any segmentation and 3D modeling.

## 2 View generation based on a linear imaging model

In our approach, assuming that a scene has two depth layers, foreground (near objects) and background (far objects) at different depths  $Z_1$  and  $Z_2$  respectively, we acquire two images, a near-focused image ( $g_1$ ) and a far-focused image ( $g_2$ ). Using foreground and background textures defined by  $f_1$  and  $f_2$  respectively, we model  $g_1$  and  $g_2$  as

$$\begin{cases} g_1(x) = f_1(x) + h_2(x) * f_2(x) \\ g_2(x) = h_1(x) * f_1(x) + f_2(x), \end{cases} \quad (1)$$

where  $h_1$  and  $h_2$  are the blurring functions (Gaussian functions) of  $f_1$  in  $g_2$  and  $f_2$  in  $g_1$ , respectively,  $x$  is a pixel coordinate and  $*$  indicates convolution. Using these images, we generate an image ( $g_v$ ) of the virtual camera that can be slightly translated and arbitrarily focused. The virtual camera image  $g_v$  is assumed to be modeled as:

$$g_v(x) = h_a(x) * f_1(x - d_a) + h_b(x) * f_2(x - d_b), \quad (2)$$

where  $h_a$  and  $h_b$  are blurring effects caused by a change of the focus depth, and  $d_a$  and  $d_b$  are the displacements caused by the translation. These effects can be calculated according to the virtual camera's position and focus depth.



Figure 1: Acquired images and generated virtual view images.

Model (1) was described in [Aizawa et al. 2000], but in this sketch we present a solution that is computationally much simpler and smarter. Fourier transform (FT) of models (1) and (2) is expressed as

$$\begin{cases} G_1(\omega) = F_1(\omega) + H_2(\omega)F_2(\omega) \\ G_2(\omega) = H_1(\omega)F_1(\omega) + F_2(\omega), \end{cases} \text{ and} \quad (3)$$

$$G_v(\omega) = H_a(\omega)e^{-j\omega d_a}F_1(\omega) + H_b(\omega)e^{-j\omega d_b}F_2(\omega), \quad (4)$$

respectively, where each function written with a large letter is the FT of the corresponding small-letter function and  $\omega$  indicates frequency. Eliminating the two unknowns  $F_1$  and  $F_2$  from Eqs. (3) and (4) yields the following equation:

$$G_v(\omega) = K_1(\omega)G_1(\omega) + K_2(\omega)G_2(\omega), \quad (5)$$

where:  $K_1(\omega) = \frac{H_a(\omega)e^{-j\omega d_a} - H_b(\omega)e^{-j\omega d_b}H_1(\omega)}{1 - H_1(\omega)H_2(\omega)}$ ,

$$K_2(\omega) = \frac{H_b(\omega)e^{-j\omega d_b} - H_a(\omega)e^{-j\omega d_a}H_2(\omega)}{1 - H_1(\omega)H_2(\omega)}.$$

Equation (5) shows that the target image can be directly generated from the two original images by  $K_1$  and  $K_2$ , which are linear space-invariant and uniquely determined by  $h_a$ ,  $h_b$ ,  $d_a$ ,  $d_b$ ,  $h_1$  and  $h_2$ .

## 3 Experimental results

Experiments were performed using real images acquired with a real camera with F-number of 5.6. The two original images are shown in Figs. 2 (a) and (b). The scene consists of near and far objects which are assumed to be located at  $Z_1 = 500$  and  $Z_2 = 700$  [mm] respectively. Figures 2 (c) and (d) show the generated virtual view images when the virtual camera is horizontally translated to left by 1.8 and to right by 2.0 [mm], respectively, while focusing on the near object. It can be seen that a different displacement was rendered for each object from the difference of the distance between lines as shown in the figures. Figures 2 (e)–(h) show the generated images with a different focusing effect when the focus depth

is changed through the range 500 to 700 [mm]. Although the model is a rough approximation, it is evident that the method can produce satisfactory results.

## References

- AIZAWA, K., KODAMA, K., AND KUBOTA, A. 2000. Producing object-based special effects by fusing multiple differently focused images. *IEEE transactions on Circuits and Systems for Video Technology, 3D special issue 10*, 2, 323–330.
- POTMESIL, M., AND CHAKRAVARTY, I. 1981. A lens and aperture camera model for synthetic image generation. *Computer Graphics 15*, 3, 297–305.

\*e-mail: kubota@hal.t.u-tokyo.ac.jp

<sup>1</sup>The functions  $H_1$  and  $H_2$  can be obtained either by image processing or calibration of the optics [Aizawa et al. 2000].

# Wegzeit - the Geometry of Relative Distance

Dietmar Offenhuber

TU Vienna / Ars Electronica Center \*



Figure 1: examples of videomappings

## Abstract

This web3d project explores how the concept of non-linear space - that is space structured by relative units - can be used in VR and architecture. It offers a dynamic view on Los Angeles' structure, radically different from usual architectural representations.

## 1 Introduction

The project presented here explores how non-linear space - space that is structured by relative units - can be used in VR and architecture. It offers a dynamic view of Los Angeles' structure that is radically different from conventional architectural representations. We usually consider space as being structured by absolute units. A meter is considered to have a constant length regardless of its position in space. However, in our daily life we often use units that are relative in nature: we measure space in minutes, costs or memories. The magnitude of such parameters might be different in different locations in space. The resulting spaces are often difficult to visualize-they might be incoherent, ambiguous and contradictory. The relevance of non-linear spaces has been studied in geography, physics and psychology. Surprisingly, this topic has been widely neglected in VR so far. Space is still being treated in its Newtonian sense as being absolute and unchangeable.

### 1.1 Describing the City in Relative Space

Wegzeit is also a project about Los Angeles and how it is transformed when brought to relative space. Asking someone in L.A. about the distance between two locations usually prompts a response in minutes. It seems paradoxical that in a city with such

\* e-mail: didi@aec.at

a regular, Cartesian layout, people rely on subjective parameters for their spatial decisions. But especially here, perhaps, where the influence of real space is levelled by this regularity, the impact of relative spaces becomes more strongly visible.

## 2 Implementation

The project consists of six dynamic virtual environments that propose models of how to visualize three-dimensional relative spaces. They deal with certain properties and effects caused by the nature of relative space such as the asymmetry of temporal distances.

The environments were designed in such a way that they can be used both in immersive VR and web3d contexts. All deformation models are managed by scripted behaviors that are evaluated in real-time. Most of the underlying parameters can be changed interactively.

The creation of the environments involved little modelling; space is constructed by the actual visual perception of the driver. The frontages of the streets examined are represented by long strips of texture that were created from videos taken from the moving car.

The main concern was to use real data from our urban environment instead of working with abstract data sets. One of the biggest challenges in working with non-linear space in VR is maintaining its legibility for the user. Working carefully with internal and external views of the environment, constraining the user's motion to a set of paths, and simplifying interaction are possibilities to prevent confusion in an environment without any kind of static form.

The project can be viewed at <http://futurelab.aec.at/wegzeit>

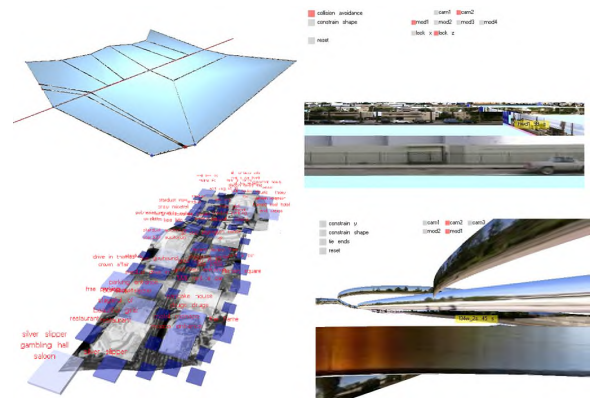


Figure 2: views of the environments



# Special Sessions – Introduction

Special Sessions encompass a wide range of topics and areas that are of current interest to most attendees. They are a way of sharing information that may not traditionally fit within the scope of other conference programs. Offered in a large audience format, these sessions inform and entertain. This year, SIG-GRAPH 2002 Special Sessions include a new way of looking at the creation of films such as *Spider-Man* and *Star Wars Episode II: Attack of the Clones* as well as an exciting look into the world of Gaming.

Of particular note this year, several animation veterans share their thoughts and experiences in a special tribute to the life and award-winning career of animator, Chuck Jones.

Special thanks to the organizers, moderators, and panelists for sharing their thoughts and outstanding achievements with us.

Jackie White, Chair

# Special Sessions – Committee

## Chair

Jackie White  
California State University - Los Angeles

## Committee

Jill Smolin  
Independent

## Special Session – Industrial Light & Magic

### Yoda and Beyond: Creating the Digital Cast of Star Wars Episode II

Computer graphics play a starring role in the production of Star Wars Episode II: Attack of the Clones. This session focuses on the creation of the digital cast of the latest prequel to the Star Wars saga. Industrial Light & Magic developed a variety of systems to make the computer-generated characters in this film stand up to the actors with whom they share the screen, both in visual quality and physical realism. These systems also made it possible for digital doubles to stand in for actors in scenes either too difficult or too dangerous to shoot practically. In an effort to match the fidelity of motion of the computer graphics characters to that of their live-action counterparts, physically based simulation was used extensively throughout the production of this film. Multi-layered clothing, skin with underlying musculoskeletal structures, and the motion of rigid bodies each played a key role in imparting a new level of physical realism into the performance of computer graphics elements. The challenge of employing this level of proceduralism is also providing methods for directing the resulting performances. In this session, we will present an overview of the pipeline and systems used to produce Episode II, with the focus of the discussion being on the specialization required to evolve technologies, deeply rooted in academic research, into effective filmmaking tools. The panel will include individuals who played key roles in the development of key digital characters for the latest prequel to Star Wars.

#### **1. Geoff Campbell, Senior Digital Model Supervisor**

The creation of a 3-D digital Yoda involved a careful transition from physically based puppet to digitally animated character.

From early screen tests and George Lucas' concerns to match Yoda's watermark performance set on *Empire Strikes Back*, this section of the talk focuses on the development of Yoda's facial performance.

#### **2. Sebastian Marino, Computer Graphics Principal Engineer**

Sebastian will discuss cloth simulation with an emphasis on the direction of the performance of the simulated garments. The depth and breadth of our pipeline will be discussed with an emphasis on how our choice of algorithms and their implementations is motivated by their application to filmmaking.

#### **3. Zoran Kacic-Alesic, Computer Graphics Principal Engineer**

Zoran Kacic-Alesic Rigid Body Dynamics is an established tool in computer animation and visual effects production at ILM. We will discuss complexity, quantity, and variety of its application in Episode II. We will also touch on some surprising ways in which artists use this tool in order to achieve a desired performance.

#### **4. James Tooley, Technical Animation Supervisor**

As digital characters play a larger role in motion pictures, there is an increasing need to enhance their appearance with simulated costumes. The specific digital costumes from *Star Wars Episode II: Attack of the Clones* will be a focus of this talk. Cloth dynamics as well as rigid body dynamics were utilized in the creation of these costumes. We will discuss methods, procedures, and problems of digitally creating costumes for a cast of digital creatures.

# The Fate of Play: Game Industry Revolutionaries Speak Out

## *Organizer*

International Game Developers Association  
[www.igda.org](http://www.igda.org)  
[info@igda.org](mailto:info@igda.org)

## *Moderator*

Jason Della Rocca  
International Game Developers Association

## *Panelists*

Raph Koster  
Sony Online Entertainment

Lorne Lanning  
Oddworld Inhabitants

Scott Miller  
3D Realms

Warren Spector  
ION Storm Austin

Will Wright  
Maxis

Prominent members of the International Game Developers Association will investigate and discuss the direction of the game industry and the impact interactive entertainment will have on our future. This panel of game industry revolutionaries will explore how game design, character development, online connectivity, business models, and social and cultural implications all weave together with advances in technology to drive the industry forward.

## **Jason Della Rocca**

Jason is the program director of the IGDA. Jason oversees the day-to-day running of the IGDA - giving particular focus to outreach efforts and member programs - working to build the sense of a unified game development community and provide a common voice for the development industry.

## **Raph Koster**

Raph is the creative director at Sony Online Entertainment where he is currently heading up the design of the highly anticipated game, Star Wars Online. Raph was the lead designer for Origin's Ultima Online.

## **Lorne Lanning**

Lorne is the president and creative director of Oddworld Inhabitants. Lorne's most recent creation was Oddworld: Munch's Oddyssey, an award winning character driven game.

## **Scott Miller**

Scott is the CEO and founder of 3D Realms. Scott was one of the driving forces behind the game industry's adoption of the shareware model of sales. Scott also helped to unleash industry icon, Duke Nukem.

## **Warren Spector**

Warren Spector is the studio director of ION Storm Austin. Warren is hard at work on the sequel to Deus Ex - one of the few games recognized for its social commentary. Warren also co-chairs the IGDA's Education Committee.

## **Will Wright**

Will is the chief designer at Maxis. Will created The Sims - the best selling computer game of all time. Among other computer simulations, Will also designed the seminal game SimCity.



## Spider-Man: Behind the Mask



Sony Pictures Imageworks takes you for a spin through the virtual world of Spider-Man. Scott Stokdyk and his team reveal how the effects, buildings, and characters were created in the computer and integrated with the live action. CG Supervisor Ken Hahn discusses the complexities involved in creating the buildings of New York City. CG Supervisor Peter Nofz explains the challenges of setting up characters for animation, and Character look Lead Greg Anderson shows us the process of look development for character lighting.

*Scott Stokdyk*  
*Visual Effects Supervisor*  
*Sony Pictures Imageworks*

Scott Stokdyk dreamed of a career creating computer generated effects for the movies, there were only a handful of people making a living at it when he started college.

Following graduation from Harvey Mudd college in 1992, Stokdyk took a job at Teradyne, Inc. in Agoura Hills, California, as an applications engineer testing microprocessors and RISC chips. Not motivated by the work and still envisioning a career in visual effects, Stokdyk added computer graphic classes at night through UCLA Extension and the American Film Institute to his Engineering Degree.

By 1994 the field of visual effects was exploding. Stokdyk landed a job at visual effects house Motions Works as a programmer. Looking for a way into effects, Stokdyk stayed late at night and taught himself to use their animation system. When two of the company's animators suddenly left, Stokdyk got his break. He started by creating animation for commercials and straight-to-video animation productions. Later in that same year he moved to Metrolight Studios as a technical director. There he worked on an AMC Gift Certificate piece, effects shots for feature film "Broken Arrow" and on Herbie Hancock's music video "Dis is Da Drum." Then in 1995 Stokdyk jumped to Digital Domain and put his visual effects career in full gear, working as a digital artist on "Titanic," "Terminator 2/3D" and as a sequence supervisor on "Fifth Element."

Stokdyk came to Imageworks in 1998. He worked as a digital artist on "Starship Troopers" (1998 Academy Award Nominee, Best Visual Effects) and "Contact." He moved up to become a computer graphics supervisor on "Godzilla." Then he

supervised 150 shots on "Stuart Little" (1999 Academy Award Nominee, Best Visual Effects). He was promoted to digital effects supervisor and received his first Academy Award Nomination for supervising the creation of the incredible disappearing man and gorilla in "Hollow Man."

Stokdyk was visual effects supervisor on "Spider-Man," where his team created the digital characters of Spider-Man and the Green Goblin, as well as a synthetic version of several locations in New York City. "No one has ever created a photoreal CG city location for a movie that's been used in this way before," said Stokdyk. "This will allow us to take the camera everywhere Spider-Man flies and see him and the world from a different vantage point. Audiences are in for something new."



*Peter Nofz*  
*CG Supervisor*  
*Sony Pictures Imageworks*

Following graduation from Technische Universitaet Munich, Germany, with the equivalent of a MS in Computer Science, Peter Nofz began his entertainment career as a Production Assistant for Gesellschaft fuer Bildende Filme (GBF) in Munich, Germany. He coordinated and supervised production of 35mm industrials for such clients as Volkswagen, BASF and Bayer.

In 1989 Peter moved to Sun Microsystems as a senior consultant/lecturer, where he taught advanced seminars in graphics programming, UI design and UNIX networking throughout Europe.

Desiring to marry his computer graphics talents with his life long love of movies, Peter returned to Northwestern University in 1990 to pursue a MFA in Radio/TV/Film. While working toward his degree, Peter was a member of the adjunct faculty for advanced 3D animation and video.

After completing his degree, Peter worked at Insight Pix, first as a Technical Director, then Senior TD/Animator. His credits included such award-winning commercials as Sega Obsidian (Golden Clio: Visual FX) and McDonald's Fish (New York Festival Gold); television productions, Warner Brothers' House of Toons, HBO's, Mr. Stitch, and feature films "Waterworld," "Executive Decision" and "Kaazam."

Peter joined Imageworks in 1998 as a Senior Technical Director on "Stuart Little" and "Hollow Man," both of which garnered Academy Award nominations for Best Visual Effects. Most recently Peter was a CC Supervisor on the international hit "Spider-Man."

*Ken Hahn*

*CG Supervisor*

*Sony Pictures Imageworks*

Ken Hahn joined Sony Pictures Imageworks in 1996 as a sequence supervisor on "Contact" and went on to work on such projects as "Godzilla," "Starship Troopers" and "Stuart Little." Ken recently supervised the construction of 3D virtual city environments for Columbia Pictures "Spider-Man." He previously led an Imageworks crew in the creation of a digital human for Columbia Pictures "Hollow Man," which was nominated for an Academy Award for Best Visual Effects.

Prior to joining Imageworks, Ken spent several years as a Technical Director at Walt Disney Feature Animation and at AT&T as a Software Engineer. Ken holds a B.S. in Computer and Systems Engineering and an M.S. in Computer Science.

*Gregory Anderson*

*Sr. Technical Director*

*Sony Pictures Imageworks*

Gregory Anderson began his career in computer graphics during his undergraduate tenure at Massachusetts Institute of Technology. As an architecture student, his work focused on using computer graphics to better visualize unbuilt spaces. His research was widely published and led to winning multiple national and international juried architecture competitions. After completing his M.S. there, he decided to expand his interests into digital effects.

He joined Sony Pictures Imageworks in 1996, working as an effects animator on "Starship Troopers," then on "Godzilla." On his next project, "Hollow Man," Greg worked as Character Look Development Lead, responsible for developing the look and lighting pipeline for the digital human and gorilla in the film.

On his most recent project, "Spider-Man," Greg reprised his role as Look Development Lead, working to design the appearance of all the digital characters in the film. This led to an opportunity to work as CG Supervisor for one of the many "Spider-Man" commercials released with the film.

# What's Up Doc?

## A Fond Remembrance of Chuck Jones

Cartoons turned a corner when Chuck Jones came to town. He transformed Walt Disney's vision to one of wit, humor and mischief. We watched as Wile E. Coyote repeatedly attempted to trap the Roadrunner, only to fall victim to his own falling anvils; Pepe Le Pew's aromatic expressions taught us everything we need to know about unsuccessful romance; Marvin Martian's Gladiator skirt, tennis shoes and romanescque helmet gave us an alternate view of aliens out to destroy the earth. And of course, Bugs Bunny's gregarious self-confidence enabled him to outwit, outsmart, and outsing any adversary. For most of the 20<sup>th</sup> Century, Chuck Jones has shaped the way we see a particular side of the world, and our art and our souls are all the better for it.

This very special Special Session unites animators from all over our industry, each of whom had a particular relationship with this fabulous man and his work. Please join us for what is sure to be an inspiring, animated session about an inspiring, animated artist.

Organizer:  
Jill Smolin  
Independent

Panelists:  
*Scott Clark*  
*Directing Animator*  
*Pixar*

"If it can be read equally with pleasure by children and adults, it will be a successful story." In 1995 I was lucky enough to have Chuck Jones relay this bit of wisdom to me personally. As a young animator-in-training my internship at his studio helped me get to Pixar a year later. It's this advice that I've tried to remember in doing my own work. His influence on my animation is, and always will be, guiding, constant, and inspiring.

*Jennifer Emberly*  
*Animation Supervisor*  
*Industrial Light & Magic*

The inspiration of Chuck Jones goes beyond any words. As a kid I found my dreams in a world of Rikki Tikki Tavi's, Bugs, Grinches, Daffy Duck's and Marvin the Martians. I grew up a bit but not enough to give up my Saturday ritual of Chuck Jones Cartoons. Many times I tested the theories of his Wile E. Coyote physics on my siblings. Searching for years to find the perfect place in an adult world to combine my imagination, passion for drawing and the desire to push the limits like Chuck always had, I went into animation. I finally got to see him at a presentation and he was every bit as animated as his Bugs Bunny....it was a thrill and an inspiration. Thanks Chuck!!

*Stephen A. Fossati*  
*PDI/DreamWorks*

...It is impossible for me to adequately relate in a few paragraphs nearly ten years of seemingly disparate lessons, ranging from the physiology of a frog ("a frog is like a quail plucked clean and slathered in grease"), to the psychology of a coward ("he thinks big and acts small"). But pressed (as editorial limitations are want to do), to boil everything down to its very

essence or even to a single word, Chuck's philosophy of and his goal in directing was *believability*.

In a very practical sense, believability to Chuck meant that since frogs blink with their lower eyelids, then so too must his frog. Little touches like this, which helped to make his characters more real were important to Chuck, but the true meaning of believability was deeper and more personal than logic or factual accuracy.

Believability was knowing that the characters you were directing were real. That somewhere Bugs Bunny, Wile Coyote and Pepe Le Pew all actually lived and breathed, triumphed and failed and then woke up the next morning to try and do it all again. Believability was knowing that that "somewhere" was inside of himself and inside everybody else who might ever watch his films.

Chuck Jones was an extraordinary director; not because he knew exactly how many frames it took for a coyote falling from a cliff to strike the canyon floor below; not because he could capture the precise attitude or expression in a drawing with a single self-confident stroke of his pencil. He was a great director because he was Chuck Jones and because Chuck Jones found humor all around him, because Chuck Jones found joy and inspiration in literature, in music, in art, in life and in himself. Because Chuck Jones truly believed that the characters he was directing were real and because Chuck Jones knew that they, like he or we, were not infallible and that they, like he or we, were simply human.

*Doug Sweetland*  
*Animator/Story Artist*  
*Pixar*

"Miles Davis said the story of jazz could be told in four words: "Louis Armstrong, Charlie Parker". In my estimation, the story of animation goes: "Walt Disney, Chuck Jones". And the four names are analogous. The first names birthed the form and the second ones stripped it down to a highly efficient, highly evolved, modern language. Miles didn't think there was a third name for jazz, and I don't think there's a third name for animation."

*Barry Weiss*  
*Senior VP of Animation Production*  
*Sony Pictures Imageworks*

I grew up on Chuck's work. Not because I had dreams of being an animator, but because I liked being a kid. And I liked being mischievous and funny, just like Bugs.

When I finally discovered animation and had the privilege of meeting Chuck, I realized why I loved this business and worshipped the man. Here he was, 80+ years old and he was still a mischievous kid. And I realized that's where great characters and animation come from, mischievous kids.

I'm still one and hope to stay that way.

# Studio – Introduction

The STUDIO provides a hands-on creative environment for transformation of all kinds. It is a place to act, a space for investigating process, and an evolving environment for transforming materials and ideas. Most importantly, it is an invitation to users to become transformed themselves. Since 1999, the STUDIO has been an opportunity to think across disciplinary boundaries and expand skill sets using the latest data-capture devices, computer applications, and output technologies.

The STUDIO is by definition a highly interactive environment. One of our goals is to transform conference attendees from passive observers to active participants. Seasoned pros sit shoulder to shoulder with novice users. Artists, technologists, scholar/educators, and students experiment with machines and tweak new software and hardware combinations. Interdisciplinary collaborations reveal new ways of working and generate unexpected outcomes.

This interactivity is made possible by a network of over 100 input devices, computers, and output devices covering some 10,000 square feet. Of particular interest is an array of 3D laser scanning and motion capture technologies that allows for the creation of original models from actual objects (and individuals!) on site. We also generate original models on site and port them to both animation programs and RP devices. High Speed LAN connections enable original experiments in educational technology, gaming, and VR. A series of 2D printers—large format color, lenticular, high-end black and white—are also featured. A “cyberfashion show” vets the latest ideas in wearable computing and techno paraphernalia. This year's theme for the STUDIO is Utopia/Dystopia. Taking its cue from the 1927 silent masterpiece "Metropolis" by Fritz Lang, the STUDIO transforms its section of the Henry B. Gonzalez Convention Center into a networked city featuring with pleasure gardens, the 2D Village, the 3D Factory, Animation Central, Fashion Boulevard, Mission Control, and the Collaboration Station.

2002 marks the first year that the STUDIO has accepted proposals through a formal juried selection process. In the Call for Participation posted on the SIGGRAPH website, an invitation was made to technologists and educators for original research that could be featured in the STUDIO. Additionally, the STUDIO, in collaboration with the Art Gallery, juried a special selection of seven “working artists.” The “working artists” appear in the Art Gallery section; the educators and technologists appear in the STUDIO section of this publication. Here is a synopsis.

*BLUIsculpt™* is an interactive virtual reality application that permits a user to freely sketch voxels inside a ten-foot cube for output as physical objects. Bill Brody, a professor at the University of Alaska Fairbanks, and his colleagues, Glenn G. Chappell and Chris Hartman of the Arctic Region Supercomputing Center at the University of Alaska Fairbanks, have developed a unique interface for the virtual reality environment of the Immersadesk using their own programming and control protocols (BLUI—Body Language User and Interface). The most recent iteration of the program allows a user to save a file as a solid 3D model in stereo-lithographic format (.stl) thereby allowing for the rapid prototyping of models developed in virtual reality. In the words of the authors, the system connects the imagination to physical form.



A tech report on the *Cinema 4D XL 3D* animation, modeling, and rendering application is given by DaShawn L. Hall of the University of the Arts at the Delaware College of Art and Design. Using original examples generated with the software, Hall focuses on the unique capabilities of the application for Character Animation, Rigging Techniques, Realistic Environmental Design, and Texturing Techniques. He also builds a bridge between its use as a professional film and animation production tool and its use by educators and students.

*Drawing Circle* is a collaborative structure for creative work conceived of by O. Makai Smith of Venturi Scott Brown Associates and James Stewart of Arizona State University. The project borrows the traditional structure and activities of a life drawing class to explore that structure's potential for digital media within the STUDIO at SIGGRAPH. Using a well-lit model or still life surrounded by easels, the authors' intent is to explore a merger of this time-tested convention with 3D modeling. They hope to provide a setting for the further development of participant's skills, and to promote the STUDIO as a place for active learning and group investigation.

A report by educators and digital media specialists Pat Johnson, Donald Wass, and Colette Stemple offers an array of projects and strategies collectively titled *Technology Curriculum Exchange*. Presenting both in the context of a panel discussion and in hands-on workshops on the STUDIO floor, the three educators share their experiences with donated technology, how they integrate digital media with traditional curricula, and how they evaluate its effectiveness in K-12 settings. Rubrics, evaluations, proofs and positive responses are shared. Methodologies, resources and experiences using 3D to enrich existing disciplines are described. Resources for obtaining technology and establishing networks to support educators are discussed.

*Collaborative Frameworks* is a concept proposal for a system that will create an archive of the activities of the STUDIO at SIGGRAPH 2002. As proposed by artist and anthropologist Lily Diaz, the project allows for the display and archiving of work done on site in such a way that it can be retrieved and replayed at a later date. Additionally, in order to reflect the diversity of content presented, Diaz is working to create the archive in a collaborative manner enlisting the help of other STUDIO participants. This is a massive undertaking considering the amplitude and quantity of data that is generated by the STUDIO.

It is my hope that the addition of published papers to the range of STUDIO activities will help to not only focus and clarify the work being presented, but to serve as a strong basis for an ongoing archive of work for the future of the STUDIO at SIGGRAPH.

Come and be transformed in this year's STUDIO!

Dan Collins, Chair

# Studio – Committee

## Chair

Dan Collins  
Arizona State University/PRISM

## Committee

Pete Braccio  
Monterey Bay Aquarium Research  
Institute

Chrissy Cain-Ramirez  
Monterey Bay Aquarium Research  
Institute

Patricia Clark  
Arizona State University

Jon Cone  
Cone Editions Press, Ltd.

David Richard Nelson  
Pixar Animation Studios

Tim Quinn  
University of California, Los Angeles

Elona Van Gent  
Grand Valley State University

## BLUIsculpt™

Bill Brody, Glenn G. Chappell and Chris Hartman, Arctic Region Supercomputing Center at the University of Alaska Fairbanks. brody@arsc.edu, chappell@arsc.edu, hartman@arsc.edu

BLUIsculpt™ is an interactive virtual reality application that permits a user to freely sketch voxels inside a ten foot cube for output as physical objects. It connects the imagination to physical form. Our interface presents the user with real-time feedback in the form of surfaces. We save a file representing the surface in rapid prototyping format. Generating a solid-by-rapid prototyping completes the cycle of perception and imagination that starts in the physical world, proceeds through vision, thought, imagination and the dance of drawing to finally arrive at tangible sculpture. Our enterprise is based on the premise that symbolic representation, of which drawing is an exemplar, derives from perception, imagination and thought. We hold that these mental activities recruit and employ the same brain machinery that is fundamental to voluntary physical movement. Such is the common wisdom of generations of artists, and is well-supported by a mass of more recent studies in cognitive and evolutionary neuroscience.

### 1. Interactive Voxel Graphics

Interactive voxel graphics in virtual reality poses significant research challenges in terms of interface, file I/O, and real-time algorithms. Voxel graphics is not so new as it is the focus of a good deal of scientific visualization. Interactive voxel creation and manipulation is much more innovative. Our support for taking the product of sketching voxels into the tangible world of rapid prototyping is unique. Scientists are understandably reluctant to manipulate data. They collect, model and predict data. The scientific analogy to interactive graphics is the generation of initial conditions to some system. It is used as a method to test scientific models. We, however, are in the business of creating new imagery. In our endeavor, science is a tool and not an end. Nevertheless, we suggest there is a whole class of interactions and associated data generation scenarios that are natural to our way of working and that are also appropriate to scientific inquiry. Annotation by sketching or painting to point to and distinguish interesting and important information is very significant for science as well as art. Annotation in 3D is difficult without a good 3D interface. Our interface is three-dimensional. In particular, our method of navigation tracks exactly with the motion of the grabbing hand.



Figure 1. Two objects created in BLUIsculpt™

### 2. Drawing in VR

Everything we do in BLUIsculpt™ is based on drawing. Drawing is fundamentally gestural. We are developing tools to draw in virtual reality. This has been a four-year research project based on the premise that the body language of pointing and grasping is an appropriate focal point for the development of a human-centered user interface. We call our approach BLUI for Body Language User Interface.

The minimal act of drawing is the trace of a moving point. A minimal gesture is the motion of a single point of focus. Interesting gestures are dynamic relationships between at least two points of interest. The meaning that derives from the physical presence of a drawing comes from the character of the path taken by the point of the drawing instrument. Another component of the meaning of a drawing derives from the empathetic perception of a viewer who connects the path with their own presumptions of what they might have felt or seen or done if they made the same drawing, i.e. moved a point along the same path.

Drawing has historically been an activity that relied on the contact between a drawing instrument and a surface. In a virtual reality application such as BLUIsculpt™, however, the requirement of any surface is absent. Indeed, a significant issue in virtual drawing is how to signal the start and finish of a line absent visual and tactile feedback from contact with a surface. A very important part of the training of an artist is how to represent the world of three dimensions on a two dimensional surface. An artist does not move the point of their drawing instrument farther away to represent distant objects, but rather laboriously learns to incorporate the rigors of two-dimensional projection into their perceptual framework. Drawing with more pressure often represents closer objects. There are artists whose mastery of this kind of drawing permit them to represent the three dimensional world of mass and form and often the sense of motion and interaction between tangible forms.

Individuals who work skillfully in the three-dimensional world, individuals such as architects, sculptors, mechanics and cabinet-makers, often describe the world to others through drawing-like gestures. These gestures are made as if the hand or hands were touching and handling tangible objects. They are gestures that often incorporate imaginary objects. At other times the gestures are sensual, performed on tangible objects, representing contemplated manipulations or how it is to feel a texture. Dancers, musicians and others use gestures to represent the paths taken by bodies as well as symbolic representations of volume, pitch and the like. Our approach is to take this analysis literally by making gestures tangible.

BLUIsculpt™ enables the user to simply and naturally sketch in space. It incorporates the purity and economy of drawing on paper with the capacity to tangibly and permanently represent form. This presents a whole new set of opportunities and challenges to expressive technology. Our process of taking freely generated sketches into the realm of physical objects completes the cycle of perception and imagination that starts in the physical world, proceeds through vision, thought, imagination and the dance of drawing to finally arrive at an object you can touch. For a more complete discussion of the concepts presented here, please look at the BLUI web site, <http://www.blui.org>.

## Tech Report

### Cinema 4D XL: Advanced 3D Software for Educators and Studio Professionals

DaShawn L. Hall, MA \*

Delaware College of Art and Design

University of the Arts

#### 1) Abstract

From the Oscar winning film, *Gladiator*, to studios, such as Pacific Data Images (PDI), Cinema 4D XL has been used for background design, modeling, and pre-visualization artwork. With its user-friendly interface, powerful animation, advanced modeling tools, and high-speed rendering technology, major film and animation studios, as well as educators have utilized it. Cinema 4D XL's advanced tools, such as a VRML plug-in and various modeling systems, can be used to integrate art with other subjects in the classroom. Because of its user-friendly architecture, Cinema 4D XL is well suited for introductory courses in 3D animation and modeling. With one of the fastest radiosity rendering engines, advanced students and studio professionals will also value the application. Students at all levels will have the capability of building professional digital portfolios.

#### 2) The Benefits of Using Cinema 4D XL in the Classroom

With its user-friendly interface and special tools, educators will find Cinema 4D XL easy to learn and teach to their students. The non-specialist can introduce 3D modeling and animation into the classroom, with no additional technical hardware or experience. One of the educational features of the program is the ability to create detailed Virtual Reality environments. VR is increasingly being used in education because it gives students the opportunity to interact and become immersed within class projects. VR can be used to motivate students to create art. 3D animation and modeling can be used to bring out creativity in student artwork. Using Cinema 4D XL, students have the ability to create professional 3D artwork with little difficulty. Because 3D animation gives the artist the ability to create the unimaginable, students will be able to create whatever comes to mind. Students therefore will be able to enhance and strengthen their art skills, as well as start building a strong professional digital portfolio.



Figure 1: Big Ben Clock Tower

© 2002 DaShawn L. Hall

#### 3) The Benefits of Using Cinema 4D XL in the Studio Environment

3D artists are constantly faced with finding new ways of creating photo realistic characters and environments. Because of the complexity of realistic movements, animators take advantage of any techniques that will help simplify the process.

Radiosity, an advanced lighting technique that helps to produce photo realistic rendered images, is yet another feature that Cinema 4D XL has to offer the 3D artist. Because of this feature, the artist is able to take any idea that has been conceptualized and bring it fully to life. This software application easily integrates images that were created in any graphics program for the use of texture mapping.

NURBS and advanced polygon modeling tools are easy to learn and master. Subdivision surfaces and HyperNURBS help the 3D artist to model very organic objects that are filled with detail and realism. Unlike most other 3D programs, models can be altered with just a click of a mouse. Because of this special feature, 3D artists are able to enhance their modeling skills, without learning complicated steps.

#### 4) Summary

Cinema 4D XL is a 3D software application that is beneficial to the educator and the studio artist. For education, it can be used to integrate art with other subjects. Cinema 4DXL is easy to learn and exposes students of all ages to 3D animation, without the need of formal training. Because of its user-friendly interface and relatively gentle learning curve, it is a vehicle for creative expression and gives students the opportunity to become more fully immersed in their artwork. For the studio artist, Cinema 4D offers tools that will streamline the processes of modeling, animation, and photo realistic rendering.

#### References

- KOENIGSMARCK, A. 2002. *Maxon Cinema 4D 7: A Workshop for 2D/3D Graphic Pros.* Peachpit Press.  
WATKINS, A. 2001. *The Cinema 4D XL Handbook.* Charles Rivers Media

\* email: [mstudios@bellatlantic.net](mailto:mstudios@bellatlantic.net)



# Drawing Circle

O. Makai Smith  
Venturi Scott Brown Associates  
smith@vsba.com

James Stewart  
Arizona State University  
james.stewart@asu.edu

## Abstract

Drawing Circle borrows the traditional structure and activities of a life drawing class to explore that structure's potential for digital media. It proposes using the familiar structure of a well-lit model or still life surrounded by easels, within the Studio at SIGGRAPH. We intend to explore a merger of this time-tested convention with 3D modeling in an attempt to provide a setting for the further development of participant's skills, and to promote the Studio as a place for active learning and group investigation.

**Keywords:** art, life drawing, mimesis, praxis, pedagogy

## 1 Introduction

Learning to draw the human form has long been a beneficial, and at times, absolutely necessary skill for a visual artist to develop. It is as important now as ever considering the huge amount of creative energy and the number of digital artists actively recreating the human world in video games, CG motion pictures, and other virtual realities. Practicing together on a common subject allows participants to learn from one another and fosters a forum for discussion about the work.

## 2 Activities

Two basic categories of activity are identified: two-dimensional drawing, and three-dimensional modeling:

1 Two-Dimensional drawing: participants interested in the examination of line, color, or shade and shadow may want to use drawing on paper, or digital processes analogous to drawing such as using a tablet in a 2D drawing package.

2 Three-Dimensional modeling: participants interested in the plastic exploration of space and form may want to physically model using plasticine clay or a 3D-modeling package. Participants choosing this activity may focus on the subject's scale and proportion, or on methods such as working from curves versus combining primitives.

By alternating physical media, 2D and 3D workstations all involved in a single activity, interplay between various areas of the Studio will be encouraged.

## 3 Intentions

As a proposal for a working method, Drawing Circle must address the particular requirements of its situation, that is, how to create a conducive environment for the hands-on practice of digital art within the context of the Studio at SIGGRAPH.

We consider here an incremental approach that examines new techniques next to conventional methods. This is not a specifically technical proposal, rather an idea about a social circumstance regarding technology. It is a framework that leaves room for improvisation regarding the technology's interaction with the creative process.

Drawing Circle aims to examine mimesis, wherein ideas are brought into being through a poetic or creative act. In this act, a

dialogue occurs between the maker and the medium, through a process often understood as craft. Without examining mimesis within technology we are merely, *and only*, technologists not artists.

In the foreground is an outlook rather than a particular technology. We are suggesting that the most advanced techniques happen at the same time, in the same place, and on the same subject as the most basic, physical techniques. The specific choice of technology is a parallel but separate discussion, since drawing on paper, sculpting in clay, and squeezing a point cloud with a haptic glove are peer activities. It is our intent that any technology available at that moment be considered.

We consider the Studio at SIGGRAPH to be neither radical, nor exclusively about technology: it is about praxis, and in that, about integration. To learn about the new is not enough, nor is it alone useful. The new must be made a part of one's own acts of doing. Thus, the benefit of conventional activities set amid the fantastically new is to see where the two meet. Drawing Circle is a proposal about method, emphasizing a procedure or process for achieving an end. It arises from a desire to propose alternatives to a conventional "computer lab" atmosphere, drawing from what we know works well in a studio. The best possible outcome for the event would be a group of people engaged in an informed public discussion about a body of work produced in the Studio, left pinned-up for display. Topics could include method (selection of medium or technique), fundamentals (line, color, texture, shadow, surface, volume, composition), advanced topics (gesture, narrative, meaning, abstraction), or topics arising from participant's questions.

## 4 Conclusions

This proposal is a pedagogical inquiry of an experimental nature. It is an opportunity for the participants to choose their methods, work along side others, and have a time to share the results. In our experience, many of the Studio's participants end up spending a few hours working alone in a crowd. We propose a straightforward alternative, to be tested as a working session during the week.

## References

- BARCSAY, JENŐ (trans. PÁL ANTAL). *Anatomy for the Artist*, 15th ed. Corvina; Budapest. 1993.
- BRIDGMAN, GEORGE BRANT. *The Human Machine*. Bridgeman Publishers, Pelham, NY. 1939.
- HOLTZMAN, STEVEN R. *Digital Mantras: The Languages of Abstract and Virtual Worlds*. MIT Press, Cambridge. 1994.
- MCCULLOUGH, MALCOLM. *Digital Craft: The Practiced Digital Hand*. MIT Press; Cambridge. 1998.
- MITCHELL, WILLIAM J., MCCULLOUGH, MALCOLM. *Digital Design Media: A Handbook for Architects and Design Professionals*. Van Nostrand Reinhold, NY. 1991.

# Collaborative Frameworks: A Proposal for an Archive in the Studio

Lily Diaz-Kommonen  
University of Art and Design Helsinki, Media Lab  
diaz@uia.fi

## 1. Abstract

This presentation describes a concept proposal for a system that will create an archive of the activities of the Studio at SIGGRAPH 2002. The initial parameters that constrain the design process are that the materials of such repository must be displayed and archived during the conference itself, and they must be formatted in such a manner that facilitates their retrieval and replay at a later date. Additionally, in order to reflect the diversity of content present, the archive must be created in a collaborative manner, and with the help of other Studio participants. The amplitude of the desired coverage is yet another factor to be considered. All of these factors must be considered in the light of the quantity of data that can be potentially generated by such endeavor.

## 2. Activity Coombs<sup>1</sup>

One can visualize the proceedings of the Studio as a progressive narrative, or performance, that evolves through time and in diverse locations, or settings. The physical settings of the Studio can be visualized as Activity Coombs. They are hollow spaces progressively filled with events in the form of the different activities of the participants. With the help of participants of the Studio and student volunteers, strategically placed input devices, such as high-quality video cameras and microphones, capture as data aspects of these activities. As part of the capture operation, data is properly tagged with metadata, sent for re-display in several locations of the Studio, and archived in pre-configured directories on a server.

At zero, *n*-level, the system contains only media artifacts of the type created as part of preparatory operations. It is reasonable to assume that these artifacts are still image files that illustrate:

- The works chosen for the Art Exhibition,
- Works in the Emergent Technologies section
- Education and production stations in the Studio.

As the narrative moves forward in time, data captured is differentiated, and stored separately, as six phases, or levels, each corresponding to a period of 24 hours for each day of the conference. For the general coverage of the Studio SIGGRAPH 2002, three basic types of media artifacts are suggested:

- *t*  
from interviews and presentations.
- Time lapse video animations, with a frame capture rate of 1 frame every four seconds, and which operate in the manner of a visual clocks recording the passage of time.
- High quality audio files resulting from interviews and presentations.

## 3. Classification and metadata strategy

A system, that uses a hierarchical classification model and a Dublin core compliant taxonomy is proposed:

- Author (John Smith)
- Title of Activity (Rapid Prototyping Workshop)
- Description (as keywords)
  - Participants (Industrial designers)
  - Objects created (Honeycomb shelf design prototype)
  - Tools used (3D graphics hardware, 3D software)
  - Institutions, communities (Design University)
  - Field(s) of knowledge represented (Design research)
- Date recorded.

## 4. Collaborative Frameworks in the Studio

Creation of an archive for the Studio is not a trivial task, but rather, one that requires that a group of participants *collaborate* in concert to achieve a common objective.<sup>2</sup> In Activity Theory, three levels of collaboration efforts have been described: Coordinated, cooperative and co-constructed collaboration. Whereas coordinated efforts involve routine, repetitive work that can be performed alone, in cooperative tasks “the actions of each agent influences the actions of others, enabling a synergistic effect.” Co-constructed activity pre-supposes reflection to promote re-elaboration and reconstruction of the work practice itself.<sup>3</sup> It is to be expected that assembling the archive will include all three types of collaboration. Anticipating coordinated tasks in this type of project, is a relatively straightforward matter. Cooperative and co-constructed endeavors, however, are not obviously discerned. Special activities to reveal and promote these will have to be designed.

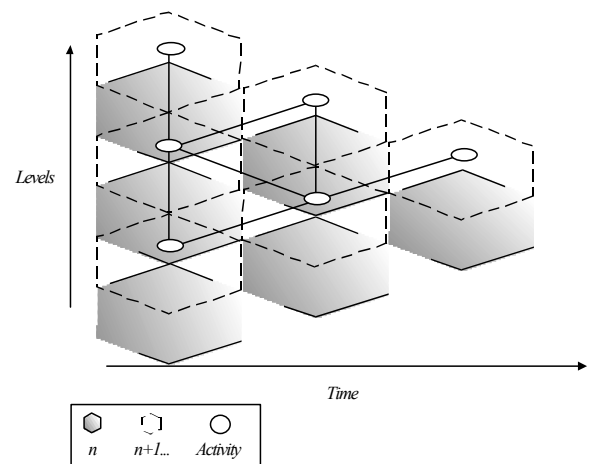


Figure 1: Studio Archive grows incrementally in time.

<sup>1</sup>. Coomb from the Greek, the hollow of a vessel. Webster Dictionary, 1913, <http://machaut.uchicago.edu/cgi-bin/WEBSTER.page.sh?PAGE=319>, (26.04.2002)

<sup>2</sup>. Engeström Y, “Coordination, Cooperation and Communication in the Courts” as cited in Barthelmes P and K. M. Anderson, “A View of Software Development Environments Based on Activity Theory” in Special Issue of CSCW on Activity Theory and the Practice of Design, Kluwer Academic Publishers. (In Press.)

<sup>3</sup>. Ibid.

# Web Graphics – Introduction

San Antonio in 2002 is the first time a web-specific program has featured in SIGGRAPH. The introduction of the 2002 Web Graphics program is designed to show SIGGRAPH attendees (animators, artists, educators, and engineers) how their talent can give them an edge in the creation of rich, interactive experiences.

Ground-breaking and inventive submissions from around the world allowed us to create a program unlike any other: brand new in scope, and yet uniquely SIGGRAPH. Content ranges from the highly experimental to the highly practical: from animation and gaming, to developing tools for collaboration across languages and borders, to new ways of navigating through information space.

The Web Graphics program has 34 sessions presented by speakers from 13 countries. They are divided into twelve categories: Images, Art & Design, Applications, Content Creation, Animation, Collaboration, Frontiers, Compression, Standards, Audio, Experimental Browsing Spaces, and Games & Communities.

Many thanks to the Web Graphics committee, whose hard work and feedback resulted in a outstanding inaugural lineup, and of course to all the contributors, whose innovative work has allowed us to present content of such quality. My gratitude to Doug Roble for allowing Web Graphics to get up to speed beside Sketches & Applications, to Dena Slothower for showing how a program should be put together, and to Vicki Caulfield for making the process run so smoothly. Thanks to the 2002 committee for their encouragement and insight creating this new program.

Simon Allardice, Chair

# Web Graphics – Committee

## Chair

Simon Allardice  
clingfish.com

## Committee

Simon Allardice  
clingfish.com

Alan Norton  
Colorado School of Mines

Colleen Case  
Schoolcraft College

Brock DeChristopher  
guitaryoga.net

Robert Reinhardt  
[theMAKERS]

Sandy Ressler  
National Institute of Standards and  
Technology; About.com

Dena Slothower  
Stanford University



## A case study in web3d film-making: Horses for Courses

This is a presentation of the web3d pipeline and production process used to deliver the multi-lingual interactive animated short film *Horses for Courses* — winner of the Web3d Roundup art prize at SIGGRAPH 2001. The film features English, French and Spanish soundtracks, different endings, mouse triggered hotspots, and manual camera controls. Photography, low-polygon modelling, and optimised texture maps, were key to publishing the film as low-bandwidth streaming media in July 2001. The pipeline developed for this project has since been used to achieve higher resolution output for offline applications.

### The idea

Our original aim was to create a short film that doubled as an interactive toy over a 56K modem connection. We also wanted to assess the suitability of web3d for delivering cost-effective cross-cultural animation. The title comes from an expression meaning a person or thing being employed for the purpose for which it is best suited. Pan was chosen as the main character since he has, like the Internet itself, been absorbed and reinterpreted by different cultures. The film was also able to demonstrate a subtle approach to interactivity where various types of interaction, such as non-linear storytelling and e-commerce, could be included without compromising the linear viewing experience. Although budget constraints meant that the pipeline had to make extensive use of off-the-shelf packages, the film's assets were kept separate from the chosen delivery format to support backwards and forward-compatibility with other media.



Pan was modelled using a real actor wearing theatrical make-up

### The web3d pipeline

All original assets, 3D models, texture maps and audio were created at broadcast quality. We chose the *b3d Studio* system to rapidly organise all assets into a library of characters, props, sets, and scenes. Original 3D models and animated sequences were created in *3D Studio Max 4.0* and exported to the *b3d* library. *B3d Studio* was used to fine-tune parameters of each scene and construct non-linear paths between scenes. The script played off the medium and included 2D visuals which helped to minimise the polygon count. Texture maps, rather than modelling, were emphasised throughout the pipeline to limit bandwidth and 3D playback requirements. Reference 360° photographs were taken of an actor wearing



All of the film's original assets were created at broadcast quality

theatrical make-up, prior to 3D modelling and animation of the main character in *Max*. Photographs were also used to plan facial morph targets created in *Max* and then referenced from within *b3d Studio* as labelled 'gestures'. The pipeline removed the need to go back to *Max* for many animation tweaks. Cameras, lighting and facial animation were controlled during the editing phase in *b3d Studio*. This also facilitated the layering of interactivity such as an alternative ending, choice of three soundtracks, as well as incidental animation and audio triggered by hotspots. Control over behaviour keys meant that it was easy to add interaction feedback cues without having to create new animation sequences separately — allowing Pan to grin at the user, for example. This approach also worked well for internationalisation. Scenes which contained language elements were replicated for ease-of-management. References to extra audio, text captions, and bitmaps that contained text were adjusted accordingly in each scene. The process was similar to DVD authoring but with far greater flexibility in how the visual elements could be manipulated at run-time.

### The streamed delivery process

Our biggest challenge was to ensure delivery over a low-bandwidth modem connection — the five minute film came in at just over 3Mb. The *b3d* delivery system allowed for assets to be independently cached on the user's computer for later reference, and this made it possible to re-use assets in multiple scenes. Using this method, Pan's head was used in a web3d banner ad campaign prior to release and the first film environment, Pan's forest, was included as part of the film's 300K loading sequence. Users were able to explore the forest while waiting for the rest of the film to download. We also wanted to use the pilot to capture feedback on the user experience. By sub-dividing the film into streaming chunks, we were able to use web logs to track the most popular interactive elements. At the last count, the clear winner was the choice of language.

Michela Ledwidge  
michela@thequality.com  
<http://thequality.com/horsesforcourses>

 thequality.com

# A Distributed Interactive Composition Tool

MÁRCIO O. COSTA

Institute of Computing  
Interdisciplinary Nucleus for  
Sound Studies  
UNICAMP, Brazil  
marcio.costa@ic.unicamp.br

JÔNATAS MANZOLLI

Interdisciplinary Nucleus for  
Sound Studies  
Music Department  
UNICAMP, Brazil  
jonatas@nics.unicamp.br

DAN SHARONI

Interdisciplinary Nucleus for  
Sound Studies  
UNICAMP, Brazil  
dsharoni@nics.unicamp.br

## Abstract

As an interactive, computerized, network oriented musical composition tool, *Rabisco* allows users to create stream of MIDI data in real-time by drawing simple sketches using a simple 2-D graphical tablet (pad). Utilizing a client-server architecture, *Rabisco* allows virtual joint compositions where several musicians interact over the network or the internet, each using a *Rabisco* client. Current applications includes Distance Learning, Interactive Music Composition, and Interaction with real world devices such as robots.

**Keywords:** java, remote application, computational music, algorithmic composition.

## 1 Introduction

The new web based musical instrument has been developed using the *Java2* development platform. The first step was to construct a Toolbox to control streams of MIDI data in real time (Costa & Manzolli, 2001). *Rabisco* system is built around a 2-D graphical tablet (pad), where one can “draw” one’s desired *sound*. Currently, Up to four different MIDI voices can be controlled by *Rabisco*. Each of this voices has a set of musical parameters like tempo, rhythm and instrument type. The system design is based on the MIDI protocol and was implemented using the platform and operating system independent Java language. Java is well suited for integration on the Internet as well. The Java standard MIDI synthesizer allows sound generation using the General MIDI Standard with or without a MIDI compatible sound card.

## 2 Musical Implementation

The X and Y axes of the pad area are linearly mapped in the range 0 to 127 integer values, that correspond to the MIDI Table values for Note and Velocity, respectively. The Note Map (X) is set for the chromatic scale but other scales such as *Chromatic, Major, Minor, Donian, Pentatonic, Hexatonic, Blues, Eolian and Mixolidian* are available. We implemented a map filter that keeps the note in “pitch class” relative to the predefined scales and octave modifiers according to the MIDI standard.

## 3 Interaction over a Network

The rich, built in, Java language network features, allow for smooth integration of *Rabisco* with the Internet. *Rabisco* was developed as two versions simultaneously: A client, which is a simple browser based applet, and a full-featured server side, standalone application for music composition.

The *Rabisco* applet version has less features, thus, is smaller for fast web-page loading, and to comply with the applet security limitations. Any Java 1.3 compatible browser can be used: (<http://www.nics.unicamp.br/~marcio/rabisco>).

The fully-featured Java application is a MIDI Server that imports several MIDI streams from clients over a network and exports the mixed audio back to clients. Several *Rabisco* clients, in different remote locations, can cooperate in a virtual musical performance

## 4 Graphic Interface

The Graphical User Interface (GUI) was implemented using Java Swing frames and event listeners not present in AWT.

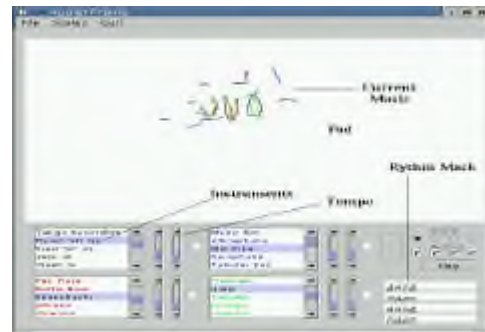


Fig.1 –  
*Rabisco*  
GUI

## 5 Conclusions and Future Work

The applet client is currently used in the Partnership in Global Learning (PGL) project. The idea was to use *Rabisco* as a fun way to enhance sound perception of children by using a ‘sound’ graphical representation. To learn more browse to:

[www.nics.unicamp.br/rabisco/pagina\\_rabisco.html](http://www.nics.unicamp.br/rabisco/pagina_rabisco.html).

The server application is being used in the Roboser and Ada projects. In this case *Rabisco* creates sound patterns to be exported as *Style Files*. See Wassermann et al, 2000 below and: [www.ini.unizh.ch/~expo/2\\_2\\_3\\_0.html](http://www.ini.unizh.ch/~expo/2_2_3_0.html).

Next steps will enhance *Rabisco* GUI to better control tempo, read and write MIDI Files, and to provide local echo while a remote performance using the MIDI server is conducted.

## References

- COSTA, M. O., AND MANZOLLI, J. 2001. Toolbox para Aplicações Musicais na Internet. In *Proceedings of SBC&M 2001*, Recife, Brazil, 125-128.
- WASSERMANN, K. C., BLANCHARD, M., BERNARDET, U., MANZOLLI, J., AND VERSCHURE, P. F. M. J. 2000. Roboser – An Autonomous Interactive Musical Composition System. In *Proceedings of ICMC 2000*, Berlin, Ed. The International Computer Music Association, 531-534.

# Adaptive Solid Texturing for Web Graphics

Bing-Yu Chen and Tomoyuki Nishita\*  
The University of Tokyo

## 1 Introduction

Solid texturing [Peachey 1985; Perlin 1985] has become a well-known computer graphics technology since it was first presented more than fifteen years ago. However, solid texturing still remains problems today, because it consumes too much time and has a very high memory requirement. Although some methods have recently been proposed to solve these problems, almost all of them need the support of hardware accelerators. Hence, these methods could not be applied to all kinds of machines, especially the low-cost ones available over the Internet. Therefore, we present a new method for procedural solid texturing in this paper. Our approach could almost render an object with procedural solid texturing in real-time using only a software solution. The basic idea of this approach is similar to the cache mechanism used for main memory control. Furthermore, to demonstrate that our approach is widely applicable we choose pure Java for its implementation, since this could not receive any benefit from the hardware and could be executed on the Internet directly.

## 2 Cache Access Storage Methodology

Basically, in order to cache the calculated texture data of a texturing function in memory, a "cache cube" is used, as other 3D-texture mapping methods require. The cache cube is empty initially, since there is no texture data has been calculated. When rendering an object with procedural solid texturing, the system checks the cache cube first to see if there is already any corresponding texture data for the drawing pixel, which is contained in a voxel. If the voxel does exist, the pixel is rendered using it, otherwise the desired texture data is calculated and stored into the cache cube as a voxel at its corresponding position. The pixel is then rendered with the calculated data.

Since the required texture data is calculated on demand, it is not necessary to generate a cache cube before implementing the raster process. Moreover, if the cache cube is pre-generated, even if we assume it is a low-resolution one containing just  $128 \times 128 \times 128$  voxels, it still costs an un-compressed file size of more than 8MB. This kind of huge data size is difficult to transmit through the narrow Internet bandwidth and requires a lot of memory to store, and it cannot even offer good quality visual effects as Fig. 1 left.

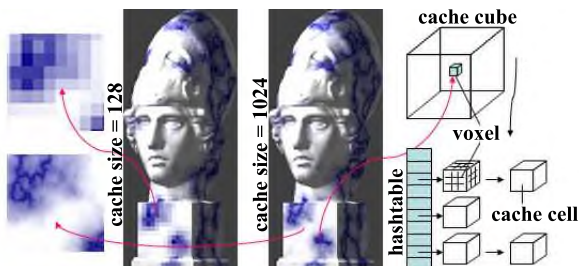


Figure 1: System diagram and results of different cache cube size.

Fortunately, when rendering an object, the visible area is only a small part of the whole shape, because the back and inside faces

is obviously not rendered. This rule is independent of the object's complexity, the number of polygons, and the calculation of the texturing function. Therefore, there is only a relatively small amount of texture data needed to be calculated and stored in the cache cube, i.e. the rest of the cache cube remains empty. To store the calculated texture data in the sparse cache cube, we separate each texture coordinate into two parts, one of which is global index, and the other is local index. Therefore, the sparse cache cube is subdivided into several cells due to the global index. Each cell is classified as either a cache cell or just as an empty cell if there is no voxel located within it. To make the retrieval of the voxel efficient, we take the size of the sparse cache cube to the power of two as the requirement for some graphics libraries such as OpenGL. Therefore, we can use some efficient algorithms, like Hashtable, to store the cache cells by using the converted global index, to get good performance for insertion and traversal, as shown in Fig. 1 right. Moreover, according to the LOD (Level-of-Detail) parameter, which is based on the well-known definition of MIP-Mapping [Williams 1983], the system could provide the most suitable levels of several sparse cache cubes for the pixel on the screen. Additionally, since there are several low-cost machines over the Internet, we also provide a mechanism to control the resolution of the cache cube automatically in accordance with the capability of the client machine.

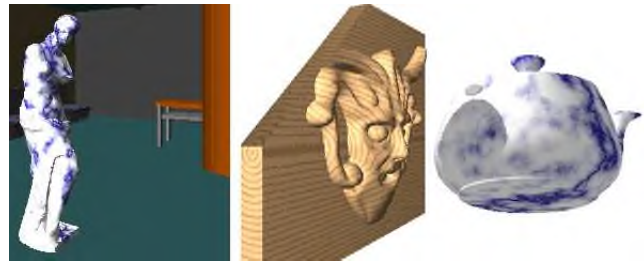


Figure 2: The results rendered with our approach.

## 3 Conclusion

In this paper, we have proposed a method to render an object with procedural solid texturing for almost all kinds of machines over the Internet<sup>1</sup>. The results are shown in Fig. 2. Although the implementation only uses pure Java, the user could also achieve an almost real-time interactive response. Our approach could be applied to arbitrary models and even their inner surface due to clipping, since it is a general solution for solid texturing, and the capability of the client machine is detected before rendering for Internet heterogeneous environment.

## References

- PEACHEY, D. R. 1985. Solid texturing of complex surfaces. *Computer Graphics (Proceedings of SIGGRAPH 85)* 19, 3, 279–286.
- PERLIN, K. 1985. An image synthesizer. *Computer Graphics (Proceedings of SIGGRAPH 85)* 19, 3, 287–296.
- WILLIAMS, L. 1983. Pyramidal parametrics. *Computer Graphics (Proceedings of SIGGRAPH 83)* 17, 3, 1–11.

\*e-mail: {robin, nis}@is.s.u-tokyo.ac.jp

<sup>1</sup><http://nis-lab.is.s.u-tokyo.ac.jp/~robin/jST/>

# Architecting a Distributed Dynamic Image Server for the Web

Alain Chesnais, Tim Beck, Rudy Ziegler  
TrueSpectra, Inc  
<http://www.truespectra.com/>

## Introduction

One of the most labor-intensive tasks in the development of web sites is the creation of derivative images for display of a visual element at multiple sizes or in multiple styles. For instance, a typical online catalog will have at least three different visual representations for each product sold: one thumbnail sized image for visual browsing, one mid sized image for viewing a product description and one large sized image for viewing product detail. The approach outlined in this presentation is to enable the web server to actively generate any derivative image from a high-resolution base image. Derivative images are then cached on the server to speed delivery of subsequent requests for the same derived content.

## Dynamic Image Serving

We call the process of enabling the server to generate derivative images on the fly "Dynamic Image Serving". In the implementation described in this presentation, commands are passed to the server via standard HTTP query strings.

Here are two examples with their associated URLs.



### Thumbnail:

<http://iris.truespectra.com/overview/phone.jpg?wid=75&cvt=jpeg>



### Medium:

<http://iris.truespectra.com/overview/phone.jpg?wid=175&cvt=jpeg>

The WID parameter that is passed to the base image in the examples above controls the width of the resulting image in pixels

The overall effect of doing this is to enable a web server to become aware of image assets and automatically perform image manipulation without requiring the use of CGI scripts or back end coding.

## Dynamic Image Serving Architecture

We propose to describe an architecture for implementing Dynamic Image Serving based upon three components:

- A cache component responsible for storing results that have already been calculated
- A render component responsible for actually doing the image manipulations that are requested
- An image store component responsible for storing the high-resolution base images and delivering the relevant portions of these images to the render component.

The challenge that arises when deploying such a solution is determining how to be able to scale as demand on the server rises. Scaling can occur in three areas: number of total hits, number of render requests and number of base images to store.

The architecture that we present allows you to implement each component described above on dedicated machines. We will describe how to monitor and scale each set of machines based on the observed load. Appropriate metrics are developed and described to show how one can then scale adequately in either of three ways (total hits, render requests, number of images) and maintain an optimally performing system. Examples are taken from our experience in providing such a networked solution in a geographically distributed Dynamic Image Serving network for the past year.



# Architectural Studios Online: The “Internet Studio Network”

Alfredo Andia, Ph.D.  
Florida International University  
andia@post.harvard.edu

## Abstract

The “Internet Studio Network” is an initiative designed to create academic relationships among architectural schools to work on semester-studio projects collaborating via the Internet. Past participants included a maximum of 300 architectural students from Miami, Argentina, Chile, Ecuador and Venezuela, who collaborated in semester-long design studios via the Internet and videoconference technology during the Fall, 2001.

**CR Categories:** K.3.1 (Computers in Education): Computer Uses in Education - Computer-managed instruction; I.3.2 (Computer Graphics): Graphics Systems - Distributed/network graphics; J.6.0 (Computer-Aided Engineering): Computer-aided design (CAD)

**Keywords:** distance education, architecture, design studio, low-bandwidth and high-bandwidth collaboration

## 1 Introduction

Traditionally, architectural studios are very isolated social environments in which approximately 15 to 40 students, and 1 to 3 professors, spend between 5 to 10 hours per week in direct contact to search for design solutions. The Internet Studio initiative investigates a different type of digital architectural experience: a design community online that relies heavily on CAD and networked technology.

## 2 Low-Bandwidth Technology

The initial experiments of the “Internet Studios” initiative explored synchronous and asynchronous collaboration with low-bandwidth technology: chat web publishing, and IP videoconferencing. Due to the different types of bandwidth connection in Latin America the most popular technology used was chat and web publishing (we called “web-chat”). Participants in online reviews, via “web-chat” mode, tended to have lower time tolerance during student reviews than in traditional studios. Reviewers usually did not want to spend more than 5 minutes with a student compared to the 15 to 20 minutes in traditional settings.

Surveys expressed that student explanations via chat were more direct, articulate, and memorable to the reviewer than oral explanations. Reviewers also expressed that comments online could be written simultaneously and reviewers were able to quickly notice the similarity of their criticism regarding the student’s work without having to wait for everybody to speak. Electronic anonymity made that the comments via chat be more direct and sometimes more violent than in face-to-face studios.

A final important observation refers to the opportunities and contact time professors and reviewers look at the student work; in traditional environments the reviewers can see the student work only during the time of the pin-up. In the online environment reviewers usually become familiar with the student web pages prior to the time when the web-chat reviews occurred.

## 3 High-Bandwidth Technology

Our experience has been more limited with high-bandwidth technology. The program is part of the AMPATH initiative at FIU, which in 2001 obtained a \$25 million grant to connect National University Networks in Latin America and the Caribbean to the Internet2 server at FIU. Only 3 countries have been able to integrate to the Ampath network in the past 2 years and Schools of Architecture in those 3 countries still struggle to those connected national networks. We have experienced 3 technologies: ISDN videoconferences, IP videoconferences over Internet2 and “Access Grid.”

The behaviors over high-bandwidth are much closer to traditional studios. There is a similar time tolerance to review student projects. However, due that the interaction usually is restricted to a television monitor, the potential for distraction at both ends is bigger. Students and professors tend to last in these environments up to 1 hour, compared to the 2 to 3 hours in traditional environment. What we saw most valuable was the potential of the technology to generate social relationships if it left to unstructured conversation.

## 4 Vision

The project is envisioned to position the “Istudio” network as a main hub for architectural education, research, and urban design thinking. The belief is that the rapid process of urbanization in the next two decades will spur similar metropolitan conditions across regions. Networking professionals and academic institutions will be essential for understanding the effects of the new speed of urbanization. The goal is to attract top urban issues, students, academics, and professionals of the region to the network.

## 5 Conclusion

This experience has generated a set of observations and conclusions, for conducting Internet studios, with low-bandwidth conditions. We are improving the web page interface with web publishing, chat, and IP video broadcasting. Istudio has been migrating to digital library ibiblio.org and hopes to develop web-tools that support interaction via regular Internet. Although our experience is limited in high-bandwidth technology we believe there is a potential for supporting more unstructured level of socialization. However the main issues that remains in this subject is access to network technology and the design of interaction spaces.

## References

- ANDIA, ALFREDO. 2001. Internet Studios: Teaching Architectural Design Online in the United States and Latin America. In *Conference Abstracts and Applications of ACM SIGGRAPH 2001*, ACM Press / ACM SIGGRAPH, C. Case, Ed., Computer Graphics Annual Conference Series, ACM, 9-12.

# Banja. Flash Programming System & Game Design

Guillaume Clary  
Team cHmAn  
gomoy@teamchman.com

## Abstract

Banja is the first communal adventure game in 3D flash on the web.

## 1 Introduction : Overall presentation of the game

The game and its goal : type of game, segmentation into episodes, an evolving game world

The principles of interaction : moving a character around on an island made up of picturesque settings; the character meets people, talks to them and can perform various actions.

The evolution of the game world over time : a new episode each month with the arrival of new characters, opening of new zones, real-time life

A persistent world. Characters, plants and animals have their own life : the cycles that are triggered by the player or independent from his actions.

The use of a database : recording of objects, actions, scores. ASP technology and network engine.

The intervention of multiplayer variables : Communal actions.

Communal aspects : Mini-games : the battle of scores. Instant messenger and player's contacts list. Chat rooms: exploitation of the motion engine. Forum : creation of a community

Evolution via publicity : dynamic advertising spaces, boutiques, juke-box, cinema, events, etc.

## 2 Exposition : Architecture of the game

Architecture of the game and basic segmentation : the zones, engines and interactions.

Links between the zones and the loading engine : organization of game zones in groups.

Links between the dialogues and the dialogue engine : segmentation of dialogues into basic elements (visualization of a dialogue, exposition of elements).

The language engine : advantages of this engine: optimization of the production.

The training mode, help during the game and inventory : ergonomics and explanation of primordial functions to the player, tools placed at the player's service in the inventory

## 2.1 Example 1 : analysis of a zone

Presentation of interaction modules in a zone : collision and hit zones, etc.

"Bwipi," : the motion and collision engine. Management of perspectives and character's graphics..

Use of a shared library for the common elements (plants, rocks, animals, etc.).

Use of a "Bwipi" de-bugging tool. Visualization of actions, objects, days and multi-player variables.

## 2.2 Example 2 : analysis of dialogue editing

The delocalization of dialogues from the zones : externalization of dialogues (repetition, a single element to be changed for updates).

Exposition of dialogue engine : saving time, creating a file ".as," a table of dialogue lines to be loaded and exploitation of the loading engine.

Exposition of the traditional dialogue structure : thanks to the dialogue engine, the creation of a simple dialogue structure: "Hello, question, answer, choice"

Making dialogue more complex by adding variable tests. Entry into a zone : interaction with the loading engine and dialogue summons from the entry zone

## 3 Conclusions and Future Work

The world of Banja is not just a technical amalgam. It opens up to new markets and is adapted to other supports such as a video game, and a 3D animated cartoon in development.

# Color Aesthetics for Web Graphics Creation

Lynda Weinman  
President, lynda.com, inc.  
lynda@lynda.com

## Abstract

Color has long played an essential role in successful visual design, but the web has some distinct constraints and considerations that warrant further study by the web graphics and SIGGRAPH community. The objective of this lecture is to establish what special constraints related to color exist, share design principles and techniques to address those constraints, and critique existing examples of web design on the success or lack thereof of color communication.

## 1 Web Color Constraints

There are numerous conditions that affect web developers. Some have to do with cross platform and hardware issues, such as gamma, bit depth, and color calibration. Having an understanding of these issues helps web developers on any platform using any technology understand what the potential conflicts are towards building successful color design. The web-safe palette will be discussed, as well as its impending (if not already impended!) obsolescence.

## 2 Color Aesthetics and Design Principles

Many web developers have no training in color theory or understanding of how to create color hierarchy in a web document. Color systems of primary, secondary, tertiary, complimentary, tertiary and analogous are important reference points, but often lack practical application. Once the starting point of color systems is understood, paying attention to value and contrast in color schemes is what allows a designer to create a hierarchy of information. Since so much of the challenge in web design and development revolves around information design, color can play a critical role in ensuring that importance is paid to important information.

## 3 Techniques

Using Photoshop, the most commonly used imaging editor for web developers, the lecture will cover techniques to create swatches and color lookup tables that promote the ability to create good contrast, color latitude for many different levels of information design, and be shared among teams of developers and designers. Practical techniques will also be covered that show how to colorize web images, from photographs to type to vectors to 3D. Files for the lecture will be available at <http://www.lynda.com/siggraph02/color>.

## 4 Critiques and Examples

It's important not to simply define the problems and solutions to successful color design, but to look to examples and references. Studies will be presented that show how to work with color systems, such as analogous and split compliments, warm and cool colors, and making hierarchy with color. As well, looking to examples on the web of existing html, flash, java, svg, and vml is important to establish a vocabulary to define the success or lack thereof of color communication on a web publication.

## About the Presenter

Lynda Weinman is the co-founder of lynda.com, an education company for creative professionals. She has worked professionally in design and animation for the past 20 years, and has been teaching digital design for the past 12 years. Lynda has been a faculty member at Art Center College of Design, San Francisco Multimedia Studies, American Film Institute, UCLA, and at numerous conferences and workshops throughout the world before forming her own workshops in Ojai, California. Lynda is the author of a series of books on web design, and writes for numerous magazines, including: MacWorld, How, Adobe, Step-by-Step Graphics, Web Techniques, Dynamic Graphics, Animation, Full Motion, New Media and Diem.

# Combining Procedural, Polygonal, and Bitmap Representations using XML

Anders Henrysson and Mark Ollila  
 Norrköping Visualization and Interaction Studio  
 University of Linköping, Norrköping Campus, Sweden

Hybrid representations - that of using image and procedural methods to synthesize images. Procedural methods allow us to describe media (2D images, 3D objects etc.) with very little information and render it photorealistically. Since the procedure is run on the client (for instance a PC or a mobile phone with limited network), it makes sense to adapt the procedure to the properties of the client.

The traditional usage of procedural methods has been to use low level shading languages or specialized applications only running one procedure and thus limiting the user to one category of media. Even though a lot can be represented with procedural methods there is a need to use traditional media representations (bitmaps, polygons etc.) as well.

We have come up with a concept where we achieve all of the above using hybrid representations. We have adopted an object based media representation where an object can either be represented with a procedure or its traditional representation. To keep the application as small and flexible as possible, each procedure is implemented as a library which is loaded when needed. The media representation is written in XML to make it human readable and easy to edit. The application is thus document driven where the content of the XML document determines which libraries to be loaded. The media objects resulting from the procedures are then composited to the media representation preferred by the renderer together with the non-procedural objects. The parameters in the XML document are relative to parameters determined by the system properties (resolution, performance etc.) and thus adapting the procedures to the client. By mapping objects to individual libraries, the architecture is easy to make multi-threaded and distributed.

The input to the procedures are in text-files pointed to in the XML document and parsed by the procedure library. The procedures in the example presented here include texture synthesis techniques and Perlin noise cloud. To make the procedures adapt to the system properties, the objects are re-rendered when the resolution is changed. The bitmap object is simply scaled, but other objects are recomputed. An array of the appropriate size is created in the main application and a pointer to it is sent to the procedure along with the name of the input file and the size of the region to be rendered. The implementation is limited but shows the strength of the concept and is a work in progress.

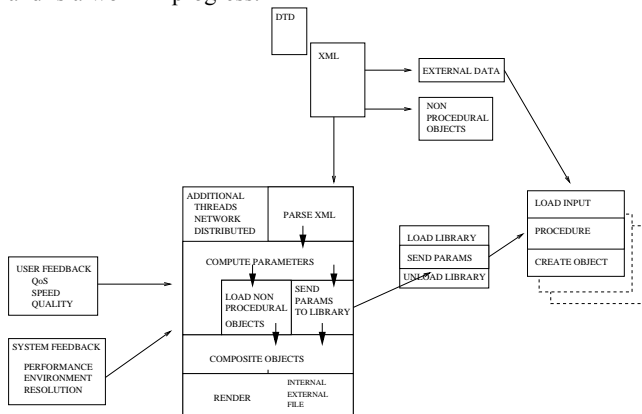


Figure 1: System design of hybrid representations.

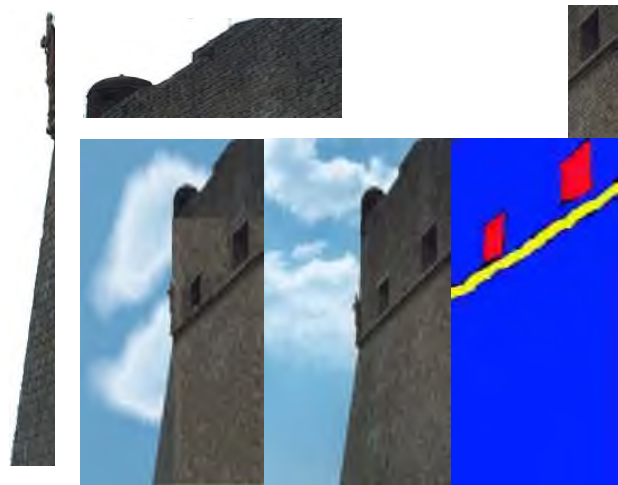


Figure 2: From left moving clockwise across the top, we have two bitmap Objects. We then have the sample texture and target image where synthesis takes place. This is followed by the original image, and the result of the hybrid rendering scheme.

### XML description:

```

<?xml version='1.0' encoding='utf-8'?>
<!DOCTYPE image SYSTEM "image.dtd">
<image size_y='550' size_x='378'>
<region size_y='100' size_x='100' pos_y='0' pos_x='0'>
<procedure type="cloud">blaise.pcf</procedure>
</region>
<region size_y='77' size_x='49' pos_y='23' pos_x='51'>
<procedure type="textureSynthesis">blaise.tsf</procedure>
</region>
<region size_y='23' size_x='51' pos_y='0' pos_x='51'>
<procedure type="bitmap">top.png</procedure>
</region>
<region size_y='60' size_x='12' pos_y='40' pos_x='40'>
<procedure type="bitmap">left.png</procedure>
</region>
</image>
    
```

### Parameter file cloud "blaise.pcf":

```

4 #number of colors to form gradient
255 102 161 201 0 #R G B offset for color one
255 102 161 201 2 #R G B offset for color two
255 224 238 249 40 #R G B offset for color three
255 159 199 225 100 #R G B offset for color four
2 #number of ellipses
50 30 48 15 #x.center y.center x.radius y.radius for ellipse 1
50 60 48 15 #x.center y.center x.radius y.radius for ellipse 2
0.7 #whispiness
    
```

### Parameterfile texture synthesis"blaise.tsf":

```

1 #has target
blaise_t.jpg #target image
1 #number of input images
blaise_in_0_100.jpg#input image 1
0 100 #y,x coords for image 1 in the target image
12 #num: num x num = # of sample points region
30 #regsize: size of region to be quilted
0.45 #overlap: size of overlap between regions
0.85 #alpha:
    
```



# Content Creation with Anark Studio

Eric Morin  
Anark Corporation  
eric.morin@anark.com

## Introduction

While setting out to create a multimedia authoring tool, the Anark development team set out to use both common design paradigms and evolving technologies. Considering current Web graphics and general multimedia trends, there is a thrust for incorporating multiple types of media into a single visual presentation. Technologies currently exist to bring the concept of pure 3D to Web graphics, but much of the difficulty in creating 3D content comes in the form of multiple and confusing workflows. What started from this thought is an investigation into the technology, use and exposure of 3D technology not only as a core component, but also as a technique to achieve new design metaphors and practices.

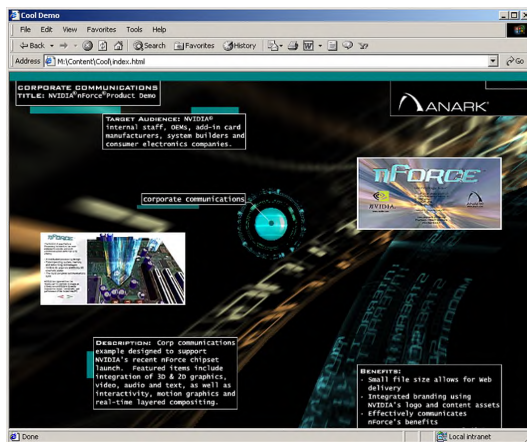
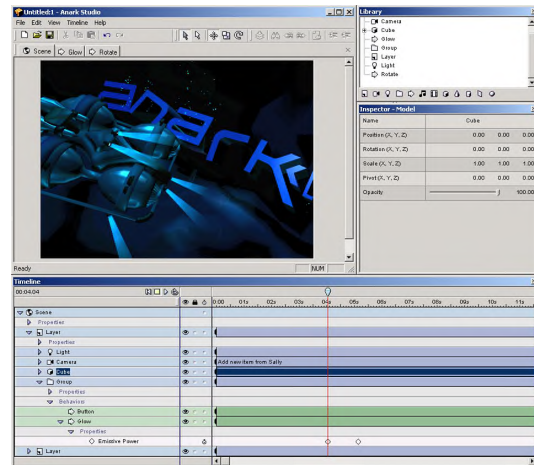
## The Development Process

The Anark development team discovered that many traditional design processes would not translate easily into this new medium. Immediately, the authoring technology presented limitless design possibilities to a content developer, but how could one instill the proper tools to express an idea? In surveying current tools and practices, many overlaps were discovered -- layers, depth, ink effects, UI, etc.

Anark's development team also began investigating traditional design tools in order to conceptually develop a seamless bridge between the user and the authoring tool. They set out to develop not just a tool with an interface complementary to the enabling technology, but also one that respects the process, practices and workflows of traditional 2D and 3D designers.

Additional experimental projects ranged from immersive environments to a more traditional layout systems. Due to the richer visuals that Anark Studio allowed, it became crucial to distinguish visually between common design and interactive elements that demanded human interaction.

Because of these issues, respecting traditional design practices and paradigms became a governing force while developing the user interface and features of Anark Studio. Using the tools that were developed, along with engaging existing design processes, the Anark development team created simple ways for a designer to cross over designs and utilize 3D and additional disparate media types.



Several situations arose during the initial development stages. For example, because the output of Anark Media could be much richer visually than a traditional interactive experience, how would the viewer interaction change?

Along with the evolution of this new visual medium came formalizing a scripting technology that adapts easily to linear and non-linear animation. The development team determined ways of combining the concepts based on object-oriented programming and current Java Script technologies. What resulted is the ability to place functional logic into a visual object/element in the same manner as one might place visual information such as color, shape, etc.

One of the problems in creating dynamic, integrated media is the merge between state-based interactivity (which is mostly non-linear) and timeline-based animation. This metaphor allows new users to grasp the scripting and programming that interactive media demands and merge them into linear, non-interactive animation and design.

## The Result

For designers interested in perusing this new medium Anark developed the software tool and the workflows and paradigms that are complementary to both traditional design institutions and advanced technologies. It is now possible for the novice and advanced designers to incorporate fully interactive, integrated media into their workflow and presentations and get immediate results.

# Creating and Implementing Photographic 3D on the web

J. Paul Nykamp\*

Diginiche

## 1 Introduction

Over the past few years, 3D on the web has been emerging as a viable format. The growth in both computing power and bandwidth has made this technically possible. 3D has many advantages over more traditional frame animations, such as QTVR object movies. These include a greater ability to display objects from any angle, and a greater ability to animate objects and have them react to user interaction. One of the challenges of creating compelling 3D, has been to make it "photorealistic". Recently a new type of 3D has emerged, which is "photographic" 3D. Photography is now being used to create both textures and geometry for 3D objects, and not just in the lab. Diginiche is doing this efficiently and commercially now, using Viewpoint technology.

## 2 Overview

Paul Nykamp of Diginiche is a leading developer of photographic 3D. He will demonstrate how he creates web 3D objects photographically from real objects using Viewpoint's 3DPhotoStudio software and other tools. Many differences between photographic 3D and traditional 3D will be outlined. He will also demonstrate how photographic 3D has been combined with other rich media formats, such as Flash, video, and Viewpoints ZoomView, to create content which is more compelling.

## 3 How Photographic 3D is unique

Photographic 3D has many differences from CG 3D, some obvious, some subtle. For instance, although the textures are generally heavier in Photographic 3D, the geometry is generally lighter. These differences will be explored.

## 4 Creating Photographic 3D

Traditional photographic skills such as studio lighting, choosing correct exposure and lenses, are combined with the latest software and hardware, such as Viewpoint's 3D Photostudio, ATI's Fire GL 8800 graphics card, Alias/Wavefront's Maya, and Adobe Photoshop to create Photographic 3D. The process has two main workflows, and these will be outlined.

## 5 Combining Photographic 3D with other rich media

Photographic 3D objects are very compelling on their own. They can be effectively combined with other rich media to create presentations which are much more informative, compelling, and interactive than the objects on their own. XML and JavaScript can be used to add animations and interactivity. Flash can be used as a texture on the 3D geometry to, for instance, bring a 3D monitor to life. Video and ZoomView can be used to compliment and enhance 3D objects. Currently live web examples, which the author created or helped create, will be used to demonstrate the possibilities:

- Ford Explorer from the Ford website
- Stanley Cup from the Hockey Hall of Fame website
- Virtual Video Host, presenting the Baird Televisor (the first commercially available tv) from the MZTV website

The technologies and processes used to create these presentations will be explained.

## 6 Web sites

[www.fordvehicles.com/suvs/explorer/3Dviews/](http://www.fordvehicles.com/suvs/explorer/3Dviews/)  
[www.legendsofhockey.net/html/silver\\_vrtro.htm](http://www.legendsofhockey.net/html/silver_vrtro.htm)  
[www.mztv.com](http://www.mztv.com) (3D Interactive gallery)  
[www.diginiche.com/explorer\\_alt](http://www.diginiche.com/explorer_alt)  
[www.diginiche.com/clothing](http://www.diginiche.com/clothing)  
[www.diginiche.com/monitor](http://www.diginiche.com/monitor)

## 7 About the Presenter

Paul Nykamp is an digital artist, photographer, and programmer. As a partner in Diginiche, a production studio based in Toronto and New York, he has helped lead the company into the forefront of cutting-edge photographic 3D and interactive virtual host technologies.

He spoke recently on the Alias/Wavefront panel at Digifest in Toronto and also gave a seminar on Viewpoint 3DPhotostudio.

He has a background in sciences from the University of Guelph, and is a graduate of the TMSI multimedia program offered jointly by York University and the Bell Centre.

---

\*email: paul@diginiche.com

# Creating Interactive 3D Web Content using AXEL

Teresa Lang

[number1@mindavenue.com](mailto:number1@mindavenue.com)

## The Technology

AXEL is a complete authoring software to create 3d interactive animation for the Internet. Developed by MindAvenue, AXEL makes it easy for designers to create interactive 3D content without scripting.

## The Objective

This virtual piano was created to allow surfers to interact freely with a piano, inside and out, using either their mouse or keyboard. Lessons were added to teach musical notation and scales.

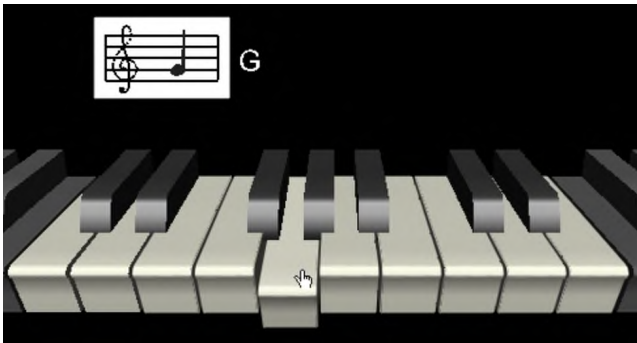


Figure 1. The Interactive Piano

## Challenges

The challenges of creating this content were similar to creating any content for the Internet: download time, playback rate and intuitive interaction. In order to achieve an acceptable sound quality and to play more than one note at a time, one challenge was to optimize the streaming of the sounds without compromising sound quality. The next challenge was to keep the visual beauty and the refresh rate high. The last challenge was to make intuitive interaction without adding pages of instructions. The target audience for this content was the general public and the interaction had to be made as easy as possible.

## Modeling and Visualization

To model the keys, we drew the profile of one white and one black key. Then we extruded the curves to create the geometry of the two keys. To copy the keys, we used instances instead of duplicates so that only one white and one black key needs to be downloaded. Then we added a white material colour to one instance and linked all the other white keys to that material; similarly all the black keys were linked to a single black material; this further reduced download time since only two materials need to be downloaded. Finally, we applied anti-aliasing only to the keys most affected rather than using it globally which would have slowed down the playback rate.

## Interaction

Interaction can be defined as something that occurs as a result of something else. In AXEL, this means creating sensors that trigger reactions.

To make the notes of the piano go down as the surfer clicks them, we created mouse sensors for each key. We linked each sensor to

a Set Orientation reaction which makes the key go down when it is clicked. We also added PlaySound reactions to each key. Every sensor has some kind of a trigger that activates it. In the case of a Mouse Sensor, the OnLeftClick trigger was used.

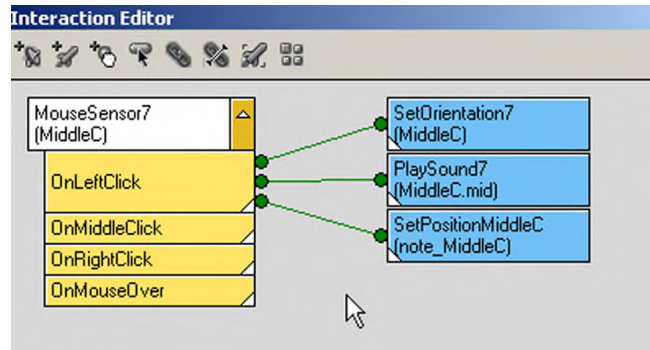


Figure 2: AXEL's Interaction Editor

## Previewing, Testing and Tweaking

An important part of creating interactive content for the Internet is to test it in browsers on the web. You can also simulate what the surfer will experience by switching to Preview mode, this displays the content as if it were downloading and being viewed in a browser on the Internet. In this example, a few things weren't quite right and were easily fixed: the key went down when clicked, but it stuck there because there was nothing to tell it to go back to its original position. To fix this, we edited the trigger link so that the reaction was triggered only for as long as the sensor was held. Another easy problem to fix was that nothing indicated to surfers that they could interact with this content. They could just as well be looking at a jpg. Mouse Sensors have a parameter called Cursor Type which we changed to Pointing Hand so that as when surfers move their mouse over a key, the cursor changes to the pointing hand, indicating that they can click the key.

## Optimizing Download Time

Sound proved to be the biggest challenge of the project. It was necessary to use wav file format in order to maintain acceptable sound quality. This resulted in stream files of between 300 and 500K. This was an issue for the target audience. Sounds were prioritized in the Stream Sequencer view to optimize this. The middle keyboard sounds were given a high priority, followed by the octaves above and below, allowing the surfer to interact with the middle keys before the rest of the content is completely downloaded.

## Learning

With the basic functionality of a piano keyboard in place, we added elements to teach the names of the notes (using Text Layers), how the notes appear on the music staff (using Set Position Reactions) and how to play a scale (using a Macromedia Flash Movie Texture).

Contributors: R&D and Creative Departments of MindAvenue.

For more information about AXEL, please visit [www.mindavenue.com](http://www.mindavenue.com)







# EVENTSCOPE: Discovering Mars with Internet-Based Virtual Environments

<http://www.eventscope.org>

Peter Coppin<sup>1</sup>, Karl Fischer<sup>1</sup>, Natalie Koch<sup>1</sup>, Dana Martinelli<sup>1</sup>, W. Ronald McCloskey<sup>1</sup>, Michael Wagner<sup>2</sup>

Tel: 1-412-268-1565, Email: [coppin@cmu.edu](mailto:coppin@cmu.edu)

**Telepresence** is experiencing a place without physically being there. Telepresence interfaces receive information from robots or sensors in distant, hard to reach places. Scientists use telepresence to explore places that are inaccessible to human beings, such as Mars. However, the technology used on such missions is so complex that the missions themselves are as inaccessible to the public as the extreme environments being studied. Subsequently, design and engineering barriers have kept this vast resource off-limits to America's classrooms despite the Internet's widespread proliferation. Existing public telepresence interfaces either do not scale well to worldwide dissemination or do not fully engage school students.

**EventScope** is a telepresence interface that enables students to discover Mars. Through EventScope, students are free to explore photorealistic virtual environments (VEs) while working through standards-based curriculum authored and downloaded to school computers via the Internet. Students then carry out activities similar to those of real mission scientists and engineers.



EventScope integrates data from robotics missions to create 3-D models that replicate remote sites in photorealistic detail. These VEs are placed within computer-based activities created by EventScope's multidisciplinary team. Classrooms can download the VEs and their associated activities into a Java3D-based user interface. EventScope was designed from the ground up to be visually engaging. Students instantly understand EventScope's interface since its metaphors are based on those of familiar video games. The use of VEs gives students the uniquely personal and powerful experience of exploration. Downloading VEs also provides freedom to experience a remote site either during a mission or after the mission is completed. Furthermore, because they are exploring a virtual world, many students can simultaneously experience the same remote location.

Educators craft telepresence curriculum with an intuitive authoring tool. It creates interactive presentations based on a sequence of "telepresence pages" that comprise a complete activity.

**EventScope** has pilot tested three iterations of software and curriculum. Pilot schools include urban, suburban and rural as well as public, private, parochial and charter schools, and represent a cross-section of needs and circumstances. EventScope bridges the digital divide between wealthier and disadvantaged school districts. It helps young people discover an interest in science by giving them the opportunity to explore real remote locations of high scientific value. Through a combination of realistic VEs and standards-based educational curriculum, students learn hard science about Mars in an entertaining, memorable way.



<sup>1</sup> STUDIO for Creative Inquiry, Carnegie Mellon U

<sup>2</sup> Robotics Institute, Carnegie Mellon University 309

# Flash MX Live: Real-time Video and Audio Delivery in Multi-user Environments

Robert Reinhardt  
[theMAKERS]

## 1. Introduction

To date, the most common forms of real-time messaging between individuals on the Internet have been text-based, including e-mail and instant messaging. Studies have suggested that Internet users overall spend more hours online with e-mail and instant messaging (IM) than with all other web browsing.<sup>1</sup> Macromedia's new technology opens up exciting possibilities for rich media applications that allow Internet users to hear and see each other instantly, using standard microphones and video devices. Although the reach of broadband connectivity is expanding into more businesses and homes, most Internet users have yet to utilize the full potential of two-way, high-speed communication available with broadband connections.

## 2. Overview

The presentation will highlight the new real-time video and audio capabilities of Macromedia's Flash Player 6 plug-in for web browsers by deconstructing a chat application that enables several listeners to hear and see a discussion between a host and any given listener in the group. This scenario is similar to a radio talk show format, but the audience can also see live video broadcast by the host and by the current "caller". The crucial issue of managing audio and video streams in a potentially crowded Internet space will also be covered. The new technology will be contextualized within an analysis of two main aspects of Internet communication: 1) the differences between one-on-one messaging and group discussion environments, and 2) the dynamics of audio and video "space" compared to that of a text-based landscape.

## 3. Deconstructing the Chat Mechanism

A portion of the presentation will be dedicated to breaking down the required technical components for two-way audio and video streaming with Macromedia Flash Player 6. Attendees will learn the process of setting up streams with client-side ActionScript and server-side scripting. Source files will be available for download after the presentation. If available, a separate session will enable users to see and work with the actual Flash documents.

## 4. Comparing Rich Media Presence to Text Messaging

After the chat application has been explained, I will discuss the differences between live audio visual chats and text-based IM, including:

- The expectation of immediate response time (and attention) in rich media communication. Text-based communication allows each participant to carefully weigh responses before sending them.

- The loss of perceived privacy in rich media communication. Text-based communication allows each participant to cloak his or her identity more easily. Moreover, with the introduction of video and audio, physical appearance, tone of voice, expression, and non-verbal communication—all components of normal human interaction—push through the anonymity of the Internet space.
- Investment in computer technology required for rich media communication. E-mail and IM are relatively inexpensive (or free) methods of communication over the Internet. Rich media communication requires an investment in high-speed connectivity and non-standard devices. Audience considerations must include some research on the penetration of web cams (or video

## 5. Considering Problems with Rich Media Communication

To conclude the presentation, consideration will be given to potential abuses of rich media connectivity among several users—including overuse of bandwidth and lack of communication management leading to user cross-talk (e.g. two or more users talking at the same time). Future modes of communication, such as instant video messaging, will also be briefly reviewed.

## 6. Web site

FlashSupport.com  
<http://chat.flashsupport.com>  
Login with any user name to channel "chat" (without the quotes).

## 7. About the Presenter

Robert Reinhardt is the lead co-author of the Flash Bible series and the forthcoming ActionScript Bible. He is also co-principal of [theMAKERS], a multimedia company based in Los Angeles.

He has developed multimedia courses for educational facilities in Canada and the United States, was a speaker at the FlashForward2000 conferences, and recently presented some of the exciting new features of Flash MX at the first-ever Canadian Flash festival, FlashintheCan 2002. With a degree in photographic arts, Robert takes a holistic approach to computer applications for the creation of provocative multimedia.

Recent projects include a Flash web site for the film *Training Day* (Warner Bros. 2001), screen graphics for Sean Penn's film *The Pledge* (Warner Bros. 2001), and on-set artwork and set design for *Gossip* (Warner Bros. 2000).

---

### Reference

1. Study: Four sites account for half of Web surfing. (2001, June 5). Retrieved April 8, 2002, from <http://www.cnn.com/2001/TECH/internet/06/05/internet.consolidation/>

# Gearation: The web3D content for children

Contributor: Fusako Nishikubo  
nishikubo@t-web.tomy.co.jp  
TOMY COMPANY,Ltd.

Collaborators: Manabu Tanaka  
g-tanaka@t-web.tomy.co.jp  
Shinta Ookino  
ookino@t-web.tomy.co.jp  
Hiroshi Ogihara  
ogihara@t-web.tomy.co.jp  
Kazuhito Ezawa  
ezawa@t-web.tomy.co.jp  
TOMY COMPANY, Ltd.

## Introduction

This web content adopts the web3D technology designed for children, especially for younger children.

We have got the idea for this project from a toy "Guru Guru" (Fig.1) of our old educational toy line, "joujou" series. What we aim at is, however, not just renewal of the existing toy. Our goal is building of a new toy concept by combining the advantages of the existing toy which help child develop their imagination and expression as well as arouse scientific curiosity, and the technology of web3D which should be able to bring what the existing toys have never achieved.(Fig.2 & 3)

## Approach

The content we are pursuing is the one that is easy for younger children to understand and give them a lot of fun. To this end we focus on the following specific approaches.

1. To provide an environment for child to play freely  
Children think and express things differently, which is something beyond our imagination. With this in mind we focus on simplicity, pursuing the development of a simple interface and achieving the minimum number of operations. In line with these efforts we have made mouse-driven input operation available.
2. To provide an environment for child to play without textual information  
We have created an environment for younger children to play comfortably by minimizing textual information and the frequent use of pictographs and small animations. We have also pursued simple interface and easy operation so that they can enjoy operation itself.
3. To create a harmony of sounds and visual images  
Each gear has its own sound. Combination of a couple of gears sounds like an ensemble in concert. As sound is a very important play element, close attention has been paid not to cause gap between sounds and visual images.
4. To bring a pure fun  
We have paid special attention to the movements and sounds of the gears so that they can fully attract child's interest and help them find a pure pleasure in making new discoveries and free expressions.

## Future Work

1. To study infant perception of 3D world in a virtual environment  
The link between children and virtual world is expected to get stronger and such trend will continue in the areas of TV, games, even on Web sites. In the field of developmental psychology a study has been conducted on how infants develop the concept of space. We think it will also be getting important to understand how they perceive the idea of virtual space and how such perception will affect their development.
2. To establish multi-user system  
As the content designed for infants does not depend on the means of language, we will be able to make the most of this characteristic. We hope to develop it into a communication tool to connect children in various countries and regions in the world and to get their messages across. We also expect that this initiative will serve as a place for experiment of our another future work just raised in the above as it could offer a great deal of valuable information including comparative data of regional differences.



Figure 1. "Guru Guru" of old educational toy line "joujou" series



Figure 2. "gearation"



Figure 3. "gearation" : The users can change the viewpoint from wherever they like to look at gears.

# Geometry Compression for ASCII Scenes

Martin Isenburg\*

University of North Carolina at Chapel Hill

## 1 Compression and VRML

The most popular way of distributing 3D content on the Web is in form of a textual representation of the scene such as VRML and its variants. The advantage of such a description is that it is very author friendly in the sense of being meaningful to the human reader. A scene represented in a textual format can be viewed, understood, and modified with any text editor (see Figure 1). Most importantly, anyone can do this, even without knowledge about the specific software package that generated the 3D content.

One disadvantage of an ASCII format is that scene files can become too large for efficient Web transmission when the scene contains many and/or detailed polygon meshes. Although the scene files are usually compressed with *gzip* compression, such general purpose coders do not come close to the compression rates that can be achieved with a dedicated geometry coder.

For image, audio, and video data an on-going standardization process has produced binary compression formats such as JPEG, GIF, and MPEG that are widely accepted. Software to read, create, save, and modify these formats is plentiful and easy to use. Efforts to develop a similar standard for compressed polygonal data have so far been without success.

Geometry compression for VRML has been an important item on the wish-list of the Web3D Consortium since 1996. It has been widely assumed that a binary format would be required to allow compressed geometry. This led to the formation of the Compressed Binary Format workgroup, which (a) created a binary format for all VRML nodes and (b) proposed new *compressed* versions of five data-heavy nodes that would only exist in binary. Despite excellent compression results, in the end the proposal was rejected. Some felt it was the sentiment against an unreadable binary format from the author-side and the reluctance to support two VRML formats from the browser company-side that influenced the decision.

## 2 Eliminating the Binary Requirement

We have recently proposed a mesh compression technique that does not require a binary file format. In [Isenburg and Snoeyink 2002] we show how to code textured polygon meshes as a sequence of ASCII symbols that compresses well with standard *gzip* compression. One benefit of this approach is that it eliminates the binary requirement. In order to add compressed geometry to VRML (or now X3D) one no longer has to wait for a binary standard. The other benefit is that complete conformance between the ASCII version and the (eventual) binary version of a VRML scene would be possible—including the compressed nodes.

However, there was a significant gap between the compression rates of binary state-of-the-art geometry coders and the proposed ASCII coder. The reason for this was that binary coders use entropy coding to squeeze the produced symbol stream into as few bits as possible. Arithmetic coding, for example, outperforms *gzip* coding because it gives optimal compression in respect to the information entropy of the symbol sequence [Moffat et al. 1998].

We have combined the advantages of arithmetic and of non-binary coding by letting the arithmetic coder produce an ASCII string of zeros and ones instead of a binary bit-stream. Subsequent *gzip* compression aggressively compresses the resulting string of zeros and ones, since it only contains two different symbols.

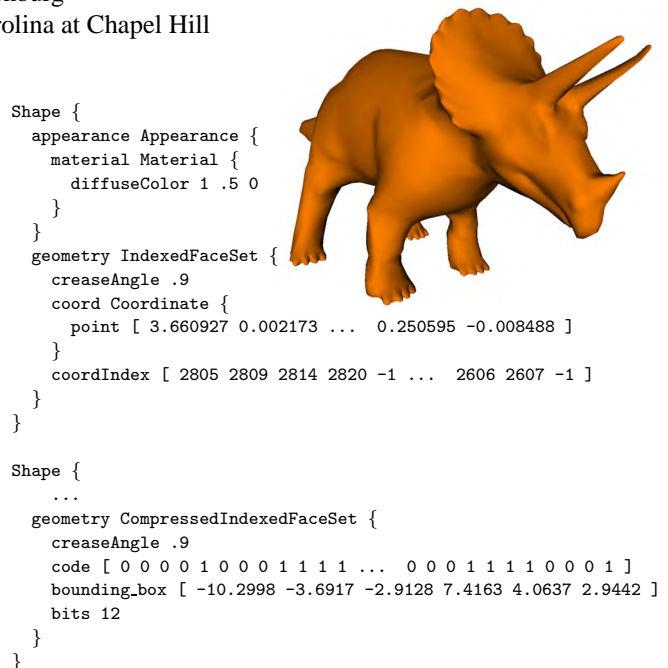


Figure 1: A simple VRML scene containing the triceratops mesh (top) and its ASCII coded version (bottom). The compressed scene can still be viewed and edited in any text editor to, for example, change the material or the crease angle. However, the size of the zipped VRML file is now only 8,392 instead of 56,307 bytes.

We have implemented an arithmetic coder that codes to and from ASCII and combined it with compression techniques for polygon mesh connectivity [Isenburg 2002] and geometry [Isenburg and Alliez 2002]. In my talk I will demonstrate that this approach is able to achieve bit-rates that are close to those of a benchmark coder.

scene	vertices	polygons	binary	plain	coded	ratio
triceratops	2832	2834	7,871	56,307	8,392	1:6.7
sandal	2636	2953	7,416	51,980	8,809	1:5.9
shark	2560	2562	6,196	49,198	5,578	1:8.8
tommygun	4171	3980	11,283	66,183	11,209	1:5.9

The table above lists compression rates for simple VRML scenes each containing a different model (like those in Figure 1). We compare the size of the zipped file containing the *plain* versus the *coded* VRML scene and report the achieved compression ratio (see our web page for an interactive demo). For comparison we also give the rates achieved by a *binary* benchmark coder [Touma and Gotsman 1998]. Quantization of vertex positions was for both coders uniformly to 12 bits of precision.

## References

- ISENBURG, M. AND ALLIEZ, P. Compressing polygon mesh geometry with parallelogram prediction. *submitted for publication*, 2002.
- ISENBURG, M. AND SNOEYINK, J. Compressing polygon meshes as compressable ASCII. In *Proceedings of Web3D Symposium'02* (Best Paper), pages 1–10, 2002.
- ISENBURG, M. Compressing polygon mesh connectivity with degree duality prediction. In *Graphics Interface'02 Conference Proceedings*, pages 161–170, 2002.
- MOFFAT, A., NEAL, R. M., AND WITTEN, I. H. Arithmetic coding revisited. In *ACM Transactions on Information Systems* 16, 3, pages 256–294, 1998.
- TOUMA, C. AND GOTSMAN, C. Triangle mesh compression. In *Graphics Interface'98 Conference Proceedings*, pages 26–34, 1998.

\*isenburg@cs.unc.edu <http://www.cs.unc.edu/~isenburg/ac>



# Integrating Multiple Narratives: *The Mirror That Changes*

Annette Weintraub  
City College of New York  
160 Convent Avenue  
New York, NY 10031  
weintraub@ccny.cuny.edu

## Introduction

This presentation describes an approach to interface and experience design in which fact and fiction are mixed in multiple narrative modes. *The Mirror That Changes* is a Flash-based Web project that integrates fact and fiction, introspection and information in several simultaneous narratives of animated text, voiceover, ambient sound and moving image. Designed to explore issues of water and sustainability from the multiple perspectives of personal use and global resource, *The Mirror* merges extensive research on water use with a sensual evocation of water so that information and symbolism mesh within a minimalist envelope.

## Shaping Content

Over the last half-century, the loosening of the conventions of narrative structure has led to more plastic kinds of storytelling and has significantly ruptured the margin between fact and fiction. New media enriches this shift in narrative form by making the movement between hard information and subjective consciousness more fluid. This renegotiated relationship between information and emotional response offers a new entry point to complex subject matter. This project mixes narrative styles, oscillating between factual and fictive elements, using a range of discursive strategies as a way of reconciling information and subjective perception.



*The Mirror That Changes* explores problems of water scarcity and sustainability through pairings of water use in domestic environments and in nature. It integrates a body of autonomous factoids and fictional threads pointing to the connection between small scale actions and large scale effects. Commonplace uses of water in domestic environments (washing clothes, bathing, cooking) find their parallel in the wider forces of nature (rain, ice and flood). Information about water use and sustainability relating to access to safe drinking water, the depletion of freshwater sources, water treated as commodity, and water pollution resulting from development and industry were organized based on this pairing structure. Elements that did not have an exact partner were paired based on a series of water metaphors, or elements of water symbolism.

Along with narrative exposition, *The Mirror* uses the visual and aural qualities of moving water to create a languid atmosphere in which overtly romantic and lyrical representations of water intersect with narratives introducing issues of scarcity, purity and equity. This conjunction of information and contemplation brings intimate scale to a global problem and mitigates the sense of information overload. *The Mirror That Changes* is a meditation on the use of water and its limits. Its interwoven fragments of fictional narrative and fact link individual action and global consequence.

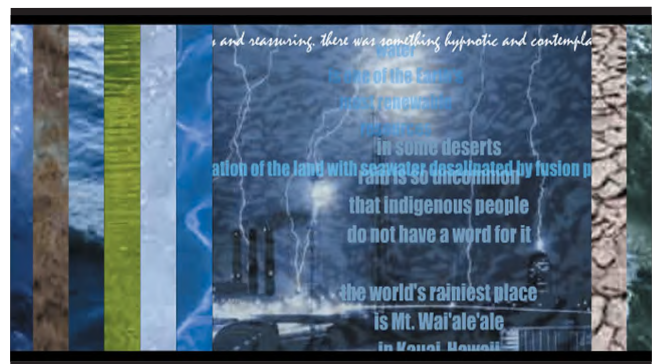
## Interface and Narrative

A primary goal of this project was to take a complex and almost numbing profusion of information and make it comprehensible yet emotionally resonant. The project's interface design emphasizes the contrast among the different qualities of water images and sounds, creating a palette of landscape element with no visible navigation elements. As the viewer moves into the piece, the complexity of elements increases, as does the density of images and information, so that the viewer is drawn into a soup of text, voiceover and moving image while being lulled by water sounds.

Providing the viewer with a dense visual flow of images and information creates a viewing experience in which one extracts bits of images and data and is able to construct alternative narratives. The whole can't be perceived at once; instead it requires a looping viewing, or alternatively, extracting a subset of the narrative from the mix as the elements advance in time. A typical strategy for Web narrative is the use of a branching structure, in which the viewer chooses her own pathway through the information. In *The Mirror*, once a particular link has been chosen, several narrative threads play simultaneously, and the larger narrative is constructed within the process of viewing and listening. This approach emphasizes the density of information, which can be sensed while at the same time gives the viewer the opportunity to follow a particular narrative line. Additionally, the mix of objective and subjective elements creates a nuanced reading that bridges the duality of the otherwise incompatible public and private realms.

## Project Design

*The Mirror That Changes* is a Flash-based work composed of numerous 'pseudomovies' which load into the primary interface animation of nine water elements. These 'pseudomovies' consist of video capture and still photo frames animated by panning, zooming and dissolve and their accompanying soundtrack of voiceover and ambient sound. Sound was a critical element in evoking the multiplicity of water, and the quality of sound was of particular concern, since severe compressions results in white noise, and not the sound of dripping, pounding, flushing, or crashing. Using Actionscript and compression in Flash, the multitrack image and sound elements preload into the interface page under the main animations. This project is contained in one HTML page.



URL: <http://www.virtualthemeworld.com/mirror>

# Interactive 3D Characters for Web-based Learning and Accessibility

Ed Sims, Ph. D.  
Vcom3D, Inc.  
eds@vcom3d.com

Dan Silverglate  
Vcom3D, Inc.  
dans@vcom3d.com

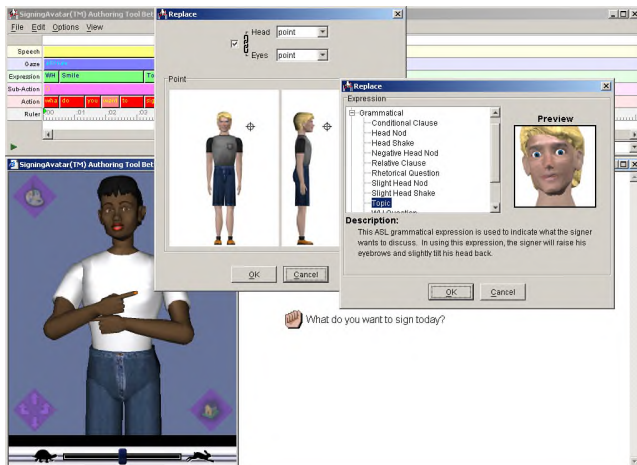
## Introduction

Interactive, life-like characters can enhance motivation, communication, and knowledge retention in computer-based learning [Lester et al. 2000]. However, until now, the development cost and computational requirements for high-quality animation have limited its widespread use.

The authors, in conjunction with a team of educators, animators and software developers at Vcom3D, Inc., are creating an Authoring Tool and Web-based Animation System that allow educators, with no previous animation experience, to create classroom and Web-based lessons with embedded animated characters. All animation is synthesized from annotated text and can be viewed using standard VRML97 and X3D browsers, thereby simplifying reuse and Web page integration.

## System Overview

Character animations are authored using an intuitive graphical user interface that combines gesture, action, facial expression, eye gaze, and synthetic speech.



### Sign Smith™ Studio Authoring Tool

Gestures can be demonstrative or emotional, or may be selected from more than 2000 signs of American Sign Language (ASL). New gestures and signs can be created within the studio using a library of hand shapes and inverse kinematic motions.

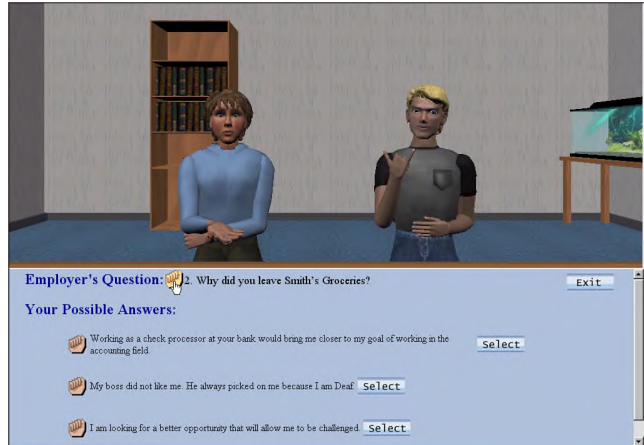
All animations and characters are developed to the Web3D Consortium's X3D and Humanoid Animation (H-Anim) Standards, allowing animations to be applied to any H-Anim character model (www.h-anim.org), and viewed in any VRML97 or X3D browser (www.web3d.org).

## Applications

Educators have used the Character Animation System to create a variety of compelling educational materials, including:

- Counseling skills training for police officers.
- An American Sign Language Dictionary.
- Interview skills training for Deaf and hearing persons.
- Science lessons on the weather, respiration, and insects.

Examples of the last two applications in the list are illustrated below.



### A SigningAvatar™ Character acts as Interpreter



### Maria assists in drawing a butterfly.

Reaction to the character animation system has been overwhelmingly positive, both by educator-authors and by learners. In one educational application, in particular, English comprehension among young Deaf learners was shown to improve from 17% to 67%.

## Reference

Lester, J., Towns, S., Callaway, C., Voerman, J., and FitzGerald, P. (2000), *Deictic and Emotive Communication in Animated Pedagogical Agents*, in *Embodied Conversational Agents*, ed. by J. Cassell et al., Massachusetts Institute of Technology, April 2000: 123-154.



## knowscape, a 3D multi-user experimental web browser

<http://knowscape.fabric.ch/knowscape/>

fabric | ch  
rue de Langallerie 6  
CH - 1003 Lausanne  
christian@fabric.ch

Christian Babski, Stéphane Carion, Patrick Keller, Christophe Guignard

### Introduction

knowscape builds new networked communities of knowledge. It explores original forms of online [user's] representation and builds new kinds of virtual world that carry on information.

knowscape is an experimental 3D browser that let users browse online content, create 3D information environment and share it [or not] with other connected users thank to its multi-user mode. Each user has a subjective 3D world, which belongs to a shared global world, and his own avatar [IP address or GSM number]. The main idea behind knowscape is to build virtual spaces and avatars made out of users \*browsing choices\*



### The technology

knowscape, which is based on our own multi-user system *rhizoreality.mu* V2.5<sup>1</sup>, demonstrates the ability for a user to modify a 3D shared world. *rhizoreality.mu* has also been used in our *La\_Fabrique*<sup>2</sup> project and will be used as well for our *electroscape.org* site that will be presented in this year's Siggraph Art Gallery. In its very last version, *rhizoreality.mu* V2.6a is able to handle arriving connection through TCP/IP as well as through UDP/IP using a circular distribution system [CDS]. Both protocols can be available at the same time. For web clients, it is possible to have a JAVA based client as well as shockwave based clients. 3D shared world can be modelled by using VRML or Shockwave 3D or even both at same time. This flexibility permits to face a large set of different needs and applications.

### knowscape functionalities

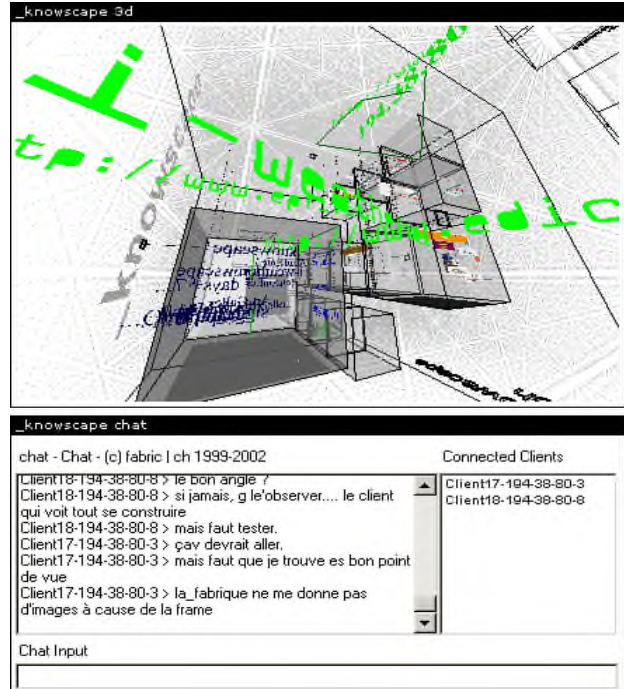
1. Persistence: knowscape can be modified globally, by adding new 3D objects permanently to the virtual world. Even if the user who added these objects leave the world, his modifications will stay visible and available for other or new users. This feature implies the management of persistence at the server level. The current state of the knowscape world is handled server side, making possible to modify which multi-user client is receiving what 3D data. It makes possible the delivering of subjective content very easily. The final user is then able to perceive the shared 3D world from its own perspective, making its experience very unique, while being still connected with other users.
2. From HTML to VRML: The conversion from HTML to VRML is a standalone service independent from the multi-user application. It is invoked by the multi-user server according to user's requests. The converter gets the HTML pages, analyses them and generates a VRML world. Then the resulting 3D scene is communicated back to the multi-user server that dispatches it to the right user. This solution avoids overloading the computer of the end-user with some low-level network tasks as well as going around security problems link to the Web context. The translator service makes first the HTTP/GET request to download the URL, then translates the content in \*true\* XML. This XML document is analysed with a DOM [Document Object Model] interface which dynamically access the structure of the page and generates a VRML world according to a set of rules based on a matching of different tags [name, attributes and values] and their 3D objects equivalents.

3. Real time avatar modification: Web site addresses visited by a user within knowscape are becoming a part of its avatar. Every new Web address is added to the user avatar (as clickable 3D text \*satellites\*). All users are aware about Web sites visited by other users just by seeing their avatars. They can decide to visit the same site on their own as well as to switch their 3D world to this particular user's view in order to see what this Web site is talking about. At any time, the user can switch back to its subjective view in order to continue to browse the Web on its side.

In its present state [conceptware], knowscape only supports/works with \*pure\* html/dhtml web pages containing no \*JavaScript redirections\* [ex. <http://www.epfl.ch>]

### Extensions

*rhizoreality.mu* and *html2vrml* are both parts of a metasytem in charge of managing heterogeneous services. This metasytem called *noosystem* has to maintain a precise image of available services at any time. Through *noosystem*, each application / service can be aware of other applications / services by declaring itself to the metasytem. It makes possible for a given application to deal with a task for which it was not programmed by finding an available service that knows how to perform this given task. From the end-user point of view, this is made seamlessly without having to invoke manually several applications. Applied to knowscape project, the conversion of HTML pages in VRML format is transparent for the end user.



### References:

- 1 Christian Babski, Patrick Keller, '*rhizoreality.mu*', Web3D symposium, 2001, ACM SIGGRAPH, pages 109-116.
- 2 Patrick Keller, Christian Babski, '*la\_fabrique: a Web3D electronic museum for the binary years*', Computer Graphics, 2000, ACM SIGGRAPH, Volume 34, Number 2, pages 66-68.

# Multilingual Flash Applications

Viswanath Parameswaran  
FastCurve Pte. Ltd.  
vnath@pacific.net.sg

## Abstract

Flash has simplified deployment of content on multiple platforms and browsers. Now with the ability to handle Unicode, it takes web applications into the next level. We will be looking at ways for creating multilingual content and a sample multilingual application using FlashMX.

**CR Categories:** H.3.5 Online Information Services, H.5.1 Multimedia Information Systems, K.8.1 Application Packages

**Keywords:** Web-based services, Graphics

## 1 Introduction

Currently Character encoding schemes, Browsers and Platforms abound. For the application developer this becomes a nightmare. A single delivery medium, which is cross platform and cross browser with support for Unicode is what Flash MX offers.

Multilingual capabilities of Flash boil down to two important aspects of manipulating text strings within Flash. The ability to bring in and to display different Character Sets.

This happens in two distinct phases. First when the movie is being created in the authoring environment and second when it is packaged as an swf file. In the authoring environment Flash depends on the Operating System to map character sets and to embed corresponding font information.

Once an swf movie, text strings are encoded in UTF-8. The swf movie plays within a Flash 6 player that can store characters both in UTF-8 and UTF-16 format. This enables us to bring in Unicode text dynamically

## 2 Single Byte Content

Russian, Greek, Turkish etc can be displayed using Single Byte encoding schemes. These characters are found in the upper ASCII range of the ASCII character set.

Usually content is brought into Flash either by

1. Cutting & pasting from a softcopy or
2. Dynamically loading from an external source or
3. By directly keying in the characters.

In the case of characters from different languages it would mean, different character sets and fonts. Since the Flash application environment depends on the Operating System, first we will have to change the language selection for the Operating System, before we cut and paste the content. In windows it would mean changing the system locale to the corresponding language and in Mac it would mean choosing the appropriate language.

Once the text is brought into Flash, you have to use the corresponding language fonts to display correctly. For static text fields, these fonts will be embedded.

## 3 Double Byte Content

The writing systems of the Far East use ideographs instead of alphabets. Each ideograph corresponds to a word. Chinese, Korean and Japanese languages use a minimum of 3000-15000 such ideographs.

Single Byte Character sets are totally inadequate to represent them. This resulted in the creation of 16 bit or Double Byte Character Sets (DBCS).

Just as we brought in Single Byte characters, we can change the system locale or languages selected in the Operating System to Chinese, Japanese or Korean and bring in these characters.

You can choose Unicode fonts or fonts in the native encoding to display these characters. Usually most of the Asian language fonts have an International name and a localized name in the local language. So you would find these names instead of the international name, once you switch to the corresponding system locale.

## 4 Dynamic Text

Dynamic text in Flash allows us to display multiple languages simultaneously. The external source should be encoded in Unicode (UTF-8 or UTF-16). Since Flash doesn't embed the entire Unicode fonts, it is better not to embed any fonts while showing multiple languages.

## 5 Flash Applications

Let us look at a sample application using these multilingual features.

### 5.1 Multilingual Dictionary

A sample Illustrated dictionary (English-Russian-Chinese Dictionary) created using Flash.

## 6 Future

Most of the display issues are due to the inability to embed specific ranges of Unicode fonts from Flash. This makes it difficult to embed fonts for Double Byte Characters or multiple Single Byte Characters together.

We can be looking at coming versions of Flash as well as third party developers to solve these issues in the future. To be a robust platform for multilingual application development Flash still has to go a long way, but for now even the small step in FlashMX is a big stride in the right direction.



# Pseudo-3D Photo Collage

Hiroya Tanaka Masatoshi Arikawa Ryosuke Shibasaki  
{tanaka,arikawa,shiba}@csis.u-tokyo.ac.jp  
Center for Spatial Information Science, The University of Tokyo

## 1 Introduction

Pseudo-3D photo collage is a new technique for creating extensive pseudo-3D scenes on the Web. This technique enables users to create, publish, navigate and share pseudo-3D scenes by easy operations. Our basic idea comes from an artistic representation "photo collage" on 2D canvases, that is, a general method of scanning and arranging original photos. Photo collage is originally 2D and static graphic representation, while our proposed representation is pseudo-3D and interactive one. Our developed system for pseudo-3D photo collage is called STAMP (Spatio-Temporal Associations with Multiple Photographs). STAMP includes basic two components as tools: *STAMP-Maker* and *STAMP-Navigator*, for creating and navigating pseudo-3D scenes respectively.

## 2 Methods

### 2.1 Creating Pseudo-3D scenes

By using *STAMP-Maker*, users apply their own photos and make spatial-hyperlinks between them (Figure 1). Spatial-hyperlinks are defined such that the rectangle specified in the photo at one end of the link is mapped onto the corresponding rectangle specified in the photo at the other end of the link. Users draw one rectangle enclosing the object on each of the two photos (Figure 2). After drawing two corresponding rectangles, the two photos are automatically distorted and superposed by matching the corresponding rectangles (Figure 2).

Users can create several spatial-hyperlinks on one photo. Therefore, a pseudo-3D scene includes many different routes for navigations. A pseudo-3D scene is a kind of hypermedia, because making an association between two different photos is similar to attaching a start anchor and an end anchor to two HTML pages in order to create a hyperlink between them.

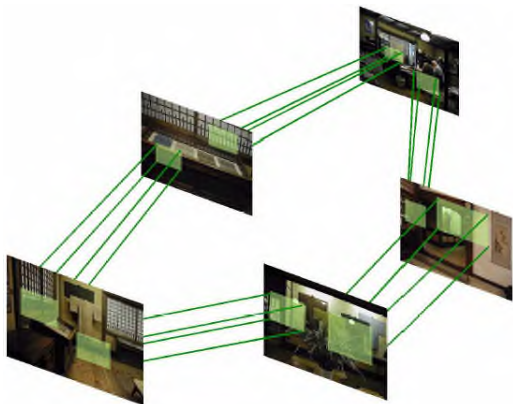


Figure1: photos and spatial-hyperlinks

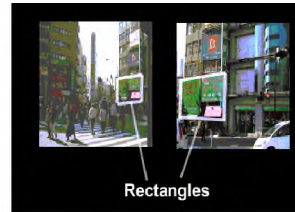


Figure 2: corresponding rectangles



Figure 3: superposed photos

### 2.2 Navigating Pseudo-3D scenes

By using *STAMP-Navigator*, users can navigate web-published pseudo-3D scenes through a simple interface. A scene consists of a key photo and the photos linked to it. The key photo is placed in the center of the display, and linked photos are placed at transformed positions (Figure 4). A scene transits to the next when users click on one of the linked photos. In a transition to the next scene, the current key photo fades out and the next key photo moves over to the center (Figure 5). This animation gives the user a sense of motion to the next position. Thus, users can freely step forward, step backward, step sideward, rotate, and zoom in/out in the pseudo-3D scenes only by clicking.



Figure 4: navigation interface

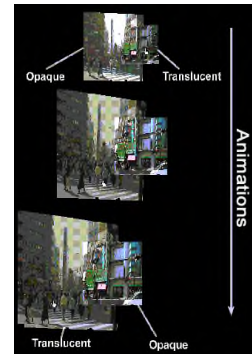


Figure 5: A scene transition

## 3 Conclusions

Both Pseudo-3D photo collage and QuickTime VR use photos to create a virtual space. QuickTime VR, however, limits itself to a panoramic scene. Pseudo-3D photo collage has an advantage of handling a depth dimension in making a virtual walkthrough. Pseudo-3D photo collage also differs from video in its ability to represent multiple paths in scene transitions, including branches, loops, reverse motions, and terminals.

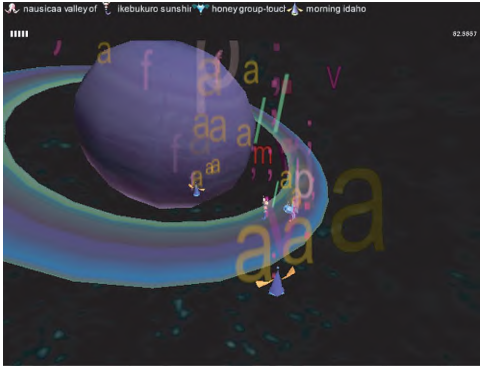
These days many people want to create and publish their original 3D contents on the Web, and therefore easy and useful tools are necessary. STAMP can be expected as a versatile tool for ordinary people to make much use of their photos in a new spatio-temporal style.

# Rhythm Engine

Hidenori Watanave  
PHOTON, Inc.  
derin@photon01.co.jp

## A Spatial Communication Tool

“Rhythm Engine” (REg) is a spatial communication tool with “Music” and “visual effects” beyond the space time. “REg” proposes an ideal way of new un-simultaneous communication to the current web-world where mainly “exchanging words” on “real time” is getting more focused.



## 1 Introduction

The world of web used to be based on un-simultaneous communication as e-mail communication indicates of its method for instance. Broad-band-ization in recent years tends to promote to pursuing of simultaneity (Real-time) connection. However, real-time communication is pursued how hard, the delay never becomes zero. Endless pursuing of convenience conceives some risks of producing a savage world. We now should re-appraise furthermore about the potential of the web, which gives us the connection with others and succession of emotion even over time and location. Through the “Rhythm Engine,” we suggest, “the session which is inherited from the past to the future using place as an agent” and “music and image communication which has rich emotion” to the Web-world.

## 2 Techniques

We are successful to open a public sphere in which a non-realtime session can be occurred by re-playing the conduct log of the player's who are logged in the system from the different time and location at same time on the “Reg” server. We studied “BBS System” and reconstruct the logic to add on “Agent log recording” and use the code measuring the key-insert timing as a score for the sound playing.

### 2.1 Structure

Play in the Rhythm Engine sphere;  
Select a character and fill in the blank; handle-name, location and e-mail address for logging in the “REg” world. See Figure 1 that some characters have been already floating freely in the “Reg public sphere.” These characters are not been operating by computer program or also not been controlling by someone who stays in front of their PC accessing at the same time=real time. Conduct log of the ex-player's character (= log records of the operation of measuring the key-insert timing) of the character has been recorded in the server and all those logs are playing back simultaneously in the “REg” space. It means that one of the characters floating in the “Rhythm Engine space” could be someone who played in ANYtime ANYwhere. There is also a possibility of being yourself of several days ago. There exist different space-time as many as characters at the same time in the “REg public sphere.”

### 2.2 Session

Encounter with someone;  
When the session begins (= encounter with any other character), the partner character hums with blowing out colorful letters. The player makes own music to communicate with others by typing keyboard. Sounds correspond to each key. See Figure 2, for example, when you type

“HELLO,” the letter objects of “H” “E” “L” “L” “O” appear overhead of the character as “visual effect” and you hear the sounds that corresponds to each key. It means that you play “the phrase” of “HELLO.” It is also able to describe that the player can leave his sound and image (could be message) to the future. All players inherit a melody from the past, and leave it for the future, and those melody lines spin forever. image (could be message) to the future.

## For The Sounds

A variety of sounds of those the gradation of phrase in a certain famous movies, which has theme of “communication with Extraterrestrials,” was analyzed are allotted to each key. Moreover, each character has different timbre of musical instruments. Percussion instrumental sound is for the octopus or synthesizer sound is for the doggy, etc.

Start a session with one person/character (duo), you continue moving then you encounter to the other character accidentally intentionally, the session will be with two people (trio). In this way, steadily the session member keeps increasing as trio, quartet and the quintet, finally up to 10 (Windows version) or 6 (Macintosh version) members possible to increase as session partners.

## 3 Future Observation

Development as a composition tool

The “Rhythm Engine” could be a composition tool to create music with yourself in the future or past or with someone on the net in the future or past wherever the person be. When the program system would be possible outsourcing the session log to score or making notes automatically indicated at the time of typing, there is a potential of new composition method opens.

Convert to the actual space

(Augmented Reality Technology)

It is possible to superpose “Rhythm Engine world” on real world using portable terminal like PDA that has the position information system such as the GPS (the Global Positioning System). In this case, the player is not to operate the character with the cursor keys but to walkout and move outside and uses the button of the portable terminal to play music and image. You can enjoy the session with anyone at anytime on the road, which always you pass, and with someone who has been passing by the place in the past.

## References

- HIDENORI “WTNV” WATANAVE (PHOTON) 2002. Concept Creation for REg & Producing. System Direction and Web master In Reg entire Project.
- TOMOYOSHI “TOM” SAITO (FROM MOST-MUSIC) 2002. System Direction and System Programming.
- HIROYA “TAGNYA” TANAKA (FROM SPACE-TIME-DESIGNS) 2002. System Direction.
- EISHIN “VOYAGE” KAWAKAMI(PHOTON) 2002. Sound Direction, Sound Programming and Texture Design.
- MARI “MARIM” ASADA (PHOTON) 2002. Art Direction and Character Design.
- ACO “H” HASHIMOTO (PHOTON) 2002. Team management

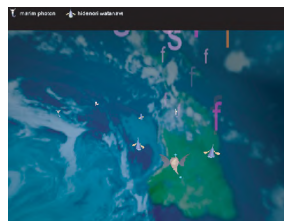


Figure 1: characters floating in the “REg public sphere.”

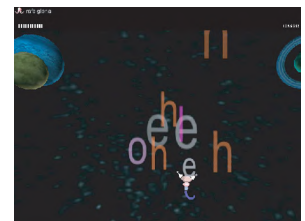


Figure 2: Colorful letters.

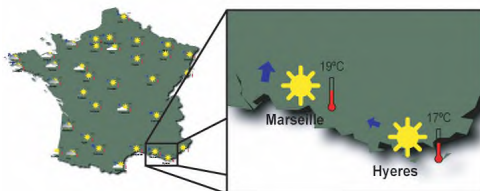
## Scalable Vector Graphics (SVG): The World Wide Web Consortium's Recommendation for High Quality Web Graphics.

Dean Jackson  
World Wide Web  
Consortium (W3C)  
<http://www.w3.org/>  
dean@w3.org



Scalable Vector Graphics (SVG) is a language for representing two-dimensional graphics. It was developed by the World Wide Web Consortium (W3C) to be the open standard format for both static and animated vector graphics on Web appliances, from desktop machines to mobile devices. The SVG 1.0 specification, whose authors include representatives from Adobe, Microsoft, Sun, Kodak, Corel, Macromedia, IBM and Apple, became a W3C Recommendation in September 2001. SVG is rapidly becoming the open standard of choice for graphics on the Web, and the many SVG implementations already in existence ensure that SVG documents can be viewed on a wide range of platforms.

The imaging model of SVG is based on industry standards, such as PostScript and PDF, with a number of Web-specific enhancements. The base features of SVG include shapes and Bézier paths, control over an object's stroke and fill, the ability to render multi-lingual text in the same manner as any other SVG shape with comprehensive font facilities, support for raster images with ICC profiles, complex fill patterns and gradients, clipping, masking, affine transformations and object transparency (using an alpha channel). More importantly, it was designed for the Web and as such includes hyperlinking into and out of the document and is encoded in XML, the core technology for current and future Web development.

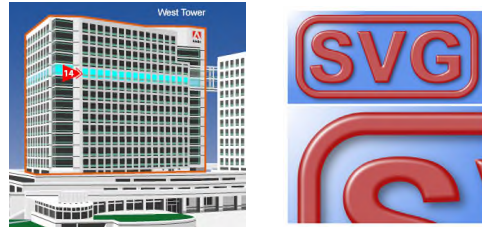


A weather map demonstrating a user ability to zoom (image courtesy ILOG)

While the majority of users currently browse the Web using a desktop machine with a fairly high resolution and color-depth, it is expected that in the near future the number of mobile browsing devices will increase significantly. SVG caters specifically to the limited resolution and processing power of these devices. The W3C has developed profiles of SVG that allow the same content to render across the wide range of Web platforms, constantly scaling to provide the best visual result.

The Web is interactive. SVG was designed from the ground up to fully integrate the user into this interactive Web environment while providing excellent quality graphics and performance. Animation is a fundamental part of SVG and is provided by SMIL, the W3C Recommendation for Multimedia. Nearly every graphical property (such as an object's shape, position, fill or stroke) can be declaratively animated in SVG, using either interpolated time-based animation, which provides the smoothest performance regardless of processing capabilities, or frame-based animation. Furthermore, the rendering tree of an SVG

document has a comprehensive object model, allowing embedded scripts in languages such as Javascript to programmatically modify any aspect of rendering, or create objects on the fly. We are beginning to see application user interfaces being designed and implemented purely in SVG. The combination of declarative animation, scripting and the complex event model make SVG images a nearly limitless interactive environment.



The interactive Adobe Building Search (left - courtesy Adobe). Lighting Effects and Drop Shadows creating a button-like appearance for SVG text (right)

A unique feature of SVG is its support for a large range of filter effects. These can be applied, independently of resolution, to the vector and raster graphics within the document. This allows the author to declaratively define effects such as drop shadows, convolution, lighting effects and image enhancement, and furthermore allow the user to control those effects through scripting and animation. With filters, SVG replaces the need for the many images on the Web that are rasterizations of text, such as page titles with drop shadows. In this case the equivalent SVG image is smaller in size, scales to provide the best representation given the requested output size, and retains the text content, ensuring it can still be selected as text (e.g. copy, cut and paste).

SVG images have benefits beyond their rendering. Accessibility is a core requirement of all W3C technologies and SVG has excellent accessibility support through metadata and the ability to provide graphical element descriptions. Also, the SVG format is fully internationalized as it uses Unicode and has built-in support for various international scripts. As a text based encoding, SVG content is available to anyone and is easily indexed by Web search engines. Many organizations are choosing SVG as their data visualization format, since it is an open format and its XML core integrates easily into their existing workflow. They can also add their own metadata to the graphics to produce a customized SVG format for their application.

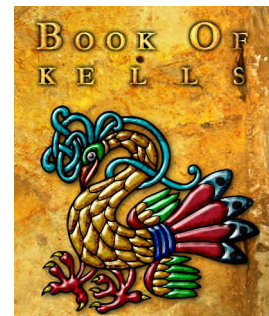
Importantly, SVG seamlessly integrates with other Web technologies – it is designed to be embedded inline within an XHTML page and all graphical elements can be styled with CSS. Given the many benefits of the format and the enthusiastic community support, the popularity of SVG as the two-dimensional Web graphics standard is set to rapidly increase.

Filter effects in SVG artwork (image courtesy of Apache Organization)

### Contributors:

Max Froumentin (W3C), Chris Lilley (W3C), Vincent Hardy (Sun), Christophe Jolif (ILOG), Jon Ferraiolo (Adobe), CSIRO Australia, Ivan Herman (W3C) and Amy van der Hiel (W3C).

See also:  
<http://www.w3.org/Graphics/SVG>  
and <http://www.w3.org/Talks/2002/svg-siggraph/>





# Scalable Visualization of Super High Resolution 3D Images for Museum Archiving

A. Basu, I. Cheng, A. Mistri, and D. Wolford  
TelePhotogenics Inc., Univ. of Alberta, Canada  
and SGI, California

[research@telephotogenics.com](mailto:research@telephotogenics.com), [www.Zoomage3D.com](http://www.Zoomage3D.com)

## Introduction and technology

In this report we describe and demonstrate our technology for creating and browsing super high resolution (SHR) & 3D digital content for a variety of applications including museum artifacts and galleries, archeology, anthropology, art design and heritage conservation. SHR 3D images and associated wireframes require large bandwidth to be transmitted in full detail. To address limitations resulting from limited bandwidth, two operations are performed: (i) bandwidth is optimally monitored using a statistical model and (ii) the quality of the 3D objects transmitted are adjusted to best fit the measured bandwidth. Regions of interest (ROIs) specified by users are stored in multiple levels of detail hierarchy. Our approach extends past systems [Martinez et al. 2000] for 2D image browsing, and supports quality of service (QoS) [Vogel et al. 1995] based retrieval. Experimental results demonstrate the feasibility of the proposed approach.

Figure 1 (left) shows a picture of the 3D scanner that can capture up to 200 Mega Pixels (approx. 5 K x 40 K x 24 bits) RGB texture & 2.5 Million or more polygons, automatically integrating 3 physically separated R, G, B channels. By comparison Minolta Vivid, Cyberware and Arius3D can capture only up to a few mega pixels and require multiple scans with different filters.



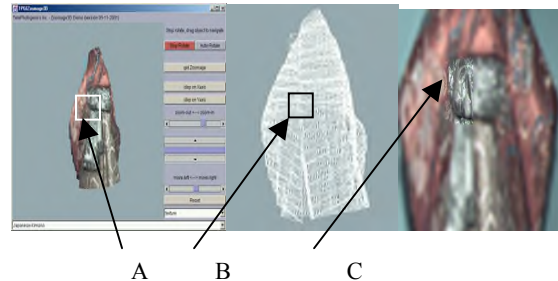
**Figure 1:** (from left to right) Zoomage3D scanning device, a tiny 3x2 sq.in. stone with writing on it, a 1 ft tall nut-cracker toy and a 18,720 polygon wireframe of the toy.

## Efficient visualization

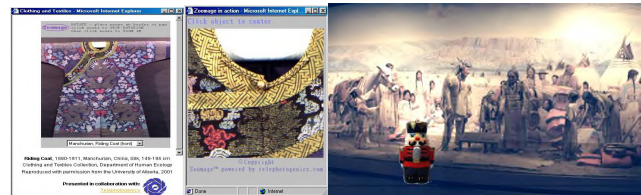
One of the problems associated with SHR 3D objects is the enormous amount of data used for object representation. This creates a major challenge for visualization and transmission. We have designed a Zoomage3D browser that can adjust the quality of a 3D object for online applications using multiple levels of detail on the wireframe of the 3D object. The texture of an object can also be stored at different levels of resolution. Depending on the available bandwidth, a simple adaptive strategy would select the resolution level and/or the number of vertices in the wireframe model to best fit the available resources. The drawback with this approach lies in the fact that the quality of the entire object is adjusted based on the available bandwidth.

A better approach would allow a user the ability to adjust quality on selected regions of interest (ROI). The approach is different from a purely hierarchical representation [Laur and Hanrahan 1991], as described in Figure 2. The ROI based approach can be

modified to a foveation [Basu and Wiebe 1998] based algorithm to make resolution change continuously from points of interest (foveae) to the periphery and allow multiple foveae to be easily integrated into one view. The technology is currently being applied to archive content at the University of Alberta (UofA), Museums & Collections (Figure 3). Beta tests are being conducted with writings on stone in Minnesota (Figure 1).



**Figure 2:** A region of interest selected on a browser (A) can be used to identify the corresponding bounding box on a wireframe (B) and associated texture to generate a high resolution region (C) that can be integrated into a lower resolution background.



**Figure 3:** (left) A Manchurian riding coat from the UofA museums; (middle) interactive online zoomed in view of neck of robe; (right) a SHR panoramic image of a gallery from the provincial museum of Alberta with a 3D object (nut-cracker toy) added in.

## Future work: Integrating 2D and 3D imaging for Museum Content

For archiving museum content it is desirable to capture panoramic scenery with a SHR camera and allow some objects in it to be visualized in 3D. Figure 3 (right) shows a gallery at the provincial museum of Alberta, with some 3D objects, such as the nut-cracker placed on the high resolution 2D scene. Viewers can both zoom into details of the 2D scene as well as view certain objects in 3D within the scene. Data acquisition for a SHR 2D-3D composite scene creation would require capture by a high resolution panoramic camera (Zoomage360) followed by 3D scan of certain objects within the scene.

A collection of demonstrations relating to the Siggraph presentation can be seen at [www.zoomage3d.com/Siggraph.htm](http://www.zoomage3d.com/Siggraph.htm).

## References

- Martinez, K. et al. 2000 Object browsing using the Internet imaging protocol. In *Proceedings of WWW9*.
- Laur, D. and Hanrahan, P. 1991 Hierarchical splatting, In *Proceedings of SIGGRAPH*.
- Vogel, A. et al. 1995 Distributed Multimedia and QoS, *IEEE Multimedia*, Summer 1995.
- Basu, A. and Wiebe, K.J. 1998 Variable resolution teleconferencing, *IEEE Transaction on Systems, Man, and Cybernetics*, 1998.



## SMIL – An Introduction

Philipp Hoschka  
World Wide Web  
Consortium (W3C)/INRIA  
<http://www.w3.org/>  
ph@w3.org



### Overview

The Synchronized Multimedia Integration Language (SMIL; pronounced "smile") enables authors to bring interactive audiovisual content to the Web. With SMIL, producing audio-visual content is easy; it does not require learning a programming language and can be done using a simple text editor.

SMIL is an XML-based language. It was developed by the W3C Synchronized Multimedia (SYMM) Working Group, a unique mix of experts from the four divergent industries (CD-ROM, Interactive Television, Web, and audio/video streaming) interested in bringing synchronized multimedia to the Web. The W3C SYMM Working Group included key players such as Compaq, CWI, INRIA, Intel, Lucent/Bell Labs, NIST, Nokia, Macromedia, Microsoft, Netscape, Oratrix, Philips and RealNetworks.

The current version of SMIL (SMIL 2.0) became a W3C Recommendation in August 2001.

### Example

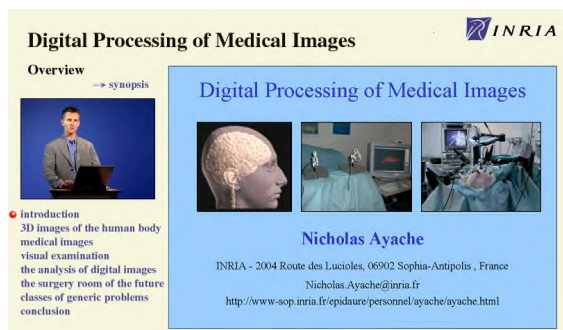


Fig 1: A SMIL Presentation on Digital Processing of Medical Images

Fig 1 shows an example SMIL presentation. A researcher was asked to come into a studio and to give a presentation of his scientific work, which was recorded on video. Then, SMIL was used to integrate and synchronize the audio and video recording with the slides used in the talk. Moreover, a table of content was added that enables viewers to navigate between different parts of the presentation (see left hand side of Fig 1, below video of speaker).

### Key Features

#### Media Objects

The core component of a SMIL presentation are individual media objects such as a video, an audio file or a set of images. All media objects are referenced via URLs rather than physically embedded into the SMIL file. For example, videos stored in a digital video library can be reused in many presentations. Different media objects in a presentation can be stored on different servers -- another benefit of using URLs.

SMIL enables authors to use clipping to extract relevant pieces of larger audio and video media objects, thus allowing editing without modifying the source file.

### Timing and Synchronization

A primary contribution of SMIL is timing and synchronization markup. SMIL elements allow the author to describe whether media objects should be played in sequence, in parallel, or are exclusive. SMIL attributes are used to refine the synchronization relation between media objects.

### Layout

Layout of media objects on a screen can be done using either a flat or a hierarchical layout model. It is also possible to assign media objects to different windows.

### Animation

SMIL animation works by changing the value of an XML attribute over time. Different types of interpolation are supported, as well as special constructs for animating colors and motion. SMIL animation is also used by SVG (Scalable Vector Graphics) to provide animation functionality.

### Content Control

SMIL content control enables authoring presentations that dynamically adapt to user preferences and to the platform to which the content is delivered. For example, authors can indicate that an audio track is available in several languages, thus increasing the potential audience of the content. Also, authors can express that a media object such as an audio track is available in different versions, each having been encoded for a different transmission bandwidth. This guarantees that presentations can be played even when only low-bandwidth access is available.

### Linking

SMIL linking markup has been inspired by HTML, and offers similar functionality. In addition, SMIL addresses timing aspects of linking. For example, it is possible to split a video into different time segments, and to associate a different hyperlink with each segment. Moreover, it is possible to control the timing behavior of the presentation that is the source of a link when the link is followed (source continues playing, pauses, stops etc.).

### Metadata

SMIL allows to associate metadata with both the full presentation and individual media objects using RDF (Resource Description Format).

### Transition Effects

Transition effects are commonly used in slideshow presentations, e.g. to fade out one slide and fade in the next when transitioning between slides. SMIL authors can use the 133 wipe transition effects defined in SMPTE 258M-1993, as well as a number of SMIL-specific effects, (push-wipe, slide-wipe, fade). SMIL transition effects can be applied not just to video and images, but to any visual media object.

### Acknowledgement:

Thanks to Dean Jackson (W3C) for motivation and help.

See also: <http://www.w3.org/AudioVideo>

# Spoiral: An online ad-lib mystery

Kazuyuki Okada  
Keio University  
Imgl/Inakage lab  
[ikki@imgl.sfc.keio.ac.jp](mailto:ikki@imgl.sfc.keio.ac.jp)

Interactive Research Group  
Keio University  
Imgl/Inakage lab  
[int@imgl.sfc.keio.ac.jp](mailto:int@imgl.sfc.keio.ac.jp)

Masa Inakage  
Keio University  
Imgl/Inakage lab  
[inakage@imgl.sfc.keio.ac.jp](mailto:inakage@imgl.sfc.keio.ac.jp)

## 1 Introduction

"Spoiral" is a multi-user online mystery game, which is implemented with web3D technology. (Fig1) Five players experience an ad-lib drama together, as if they are playing a role in a detective novel, in which the story proceeds in real time with real people.

"Spoiral" is a thrilling and intellectual game, which is epoch making in its multi-access style as a web 3D game.

## 2 Game design

The stage of the game is located in a small room shared by a little brother and sister. The world had been kept peaceful, but one day the toy robot was broken into pieces by another toy. "WHO DID IT?" - the suspect must be one of the inhabitants: the Japanese doll, the marionette, the stuffed rabbit, the baby doll or the cymbal striker. Each player operates one of these toy characters. One of the players is chosen randomly to play the suspect's role, and the rest play the detectives' roles.

The aim of the suspect is to break other toys. And the aim of the detectives is to find out who the suspect is, by collecting necessary information. But be careful, the suspect may be trying to break you! If you get broken, you lose the game and cannot keep on playing.

The detectives search the room and gather evidence of incidents. The room is divided into seven areas. The time line of the game is divided into phases that are two minutes long. Players can search only one area during one phase. When the phase ends, players can choose to move to an adjacent area or to stay in the same area. WHO and WHAT you see in WHICH PHASE and in WHICH AREA, is recorded and helps you to construct the case. Also communicating with the other characters will help you to supplement missing information.

A detective player can win when he or she reveals the truth. The suspect wins when he or she breaks all the other player characters.

All the art work and music have been designed in the "cute but scary" style called *kowaii* in Japanese.

The whole game is designed to last for approximately 30 minutes playing time.

Games with a life-risking suspense drama have just begun. We propose a new interactive content for a coming broadband era with web3D technology.

## 3 Communication and detection

To accuse the suspect, the detectives need to reveal WHO broke WHOM with WHICH WEAPON. Since the information you have got might not be enough to construct a case, getting information from other

characters is a good way to proceed. To get information from others, you go close to another character and tell him or her the AREA and the PHASE, about which you want to know (Fig.2).

When you have obtained plenty of information, you should know who had the chance to break the victim. But don't forget, the suspect is able to tell a lie about his own alibi!

This selective way of communication suggests that international playing would be easily realized.

## 4 Implementation

All the systems are programmed by Macromedia Director and the real-time 3D part is implemented with Shockwave 3D. This cost us less time and fewer human resources than implementing with c++ and DirectX. Also, it requires less hardware specification for the users and is OS independent because of the web 3D system.

A broadband connection and high performance 3D environment is strongly recommended to play "Spoiral".

<http://www.imgl.sfc.keio.ac.jp/interactive/spoiral/>



Fig1. Screen shot



Fig2. Communicating with others

# The Bridge: An Environment for Collaborative Design Learning

Yasuhiro Santo, Catherine Hu, Mamata Rao  
School of Design

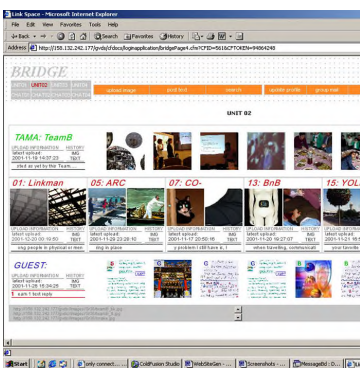
The Hong Kong Polytechnic University

[sdyasu@polyu.edu.hk](mailto:sdyasu@polyu.edu.hk), [sdcathhu@polyu.edu.hk](mailto:sdcathhu@polyu.edu.hk), [sdmamata@polyu.edu.hk](mailto:sdmamata@polyu.edu.hk)

## Bridging interactions

The idea behind the *Bridge* as a collaborative groupware arose initially with the need to find appropriate tools and environments to facilitate international design collaborations. In particular, this tool needs to support on-line collaborative activities in design learning where language is a barrier and conventional groupware fails to facilitate effective communication. Moreover, designers are inherently dependent on visual elements in all levels and phases of design development, something which current text-based groupware again fails to support.

The *Bridge* proposes to use images as the main mode of communication, with the aid of minimal textual descriptions. It aims to provide opportunities for active use of images as interactive agents for idea exchanges or collective concept development.

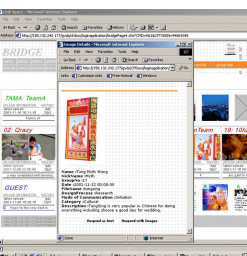


The *common area* of the *Bridge* environment mainly consists of the latest uploaded images by all collaborating participants. It is a dynamic environment that can be used to synchronously communicate and record the latest development in group communications as well as maintain a history of such communications.

Users tag simple attributes to an image while it is being uploaded. Such attributes are used both as reference anchors for the search engine, and also as loose pointers to the originator's idea for the image. Uploaded images are then displayed together with one line of optional textual information which is meant to facilitate the exchange and archival of design ideas and its progress. When the latest pictorial information from all participants is viewed together in the *Bridge*, current issues of concern or discussions or debate can easily be observed. The dynamic nature of the interface motivates participants to contribute ideas, while simultaneously promoting the social aspect of sharing within a community.

## Design Environment

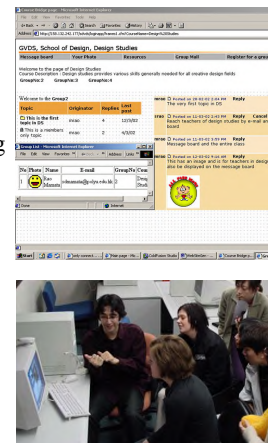
In international collaboration, language and sometimes cultural barriers, can easily discourage interactions. The *Bridge* tries to tackle this problem by using visual information as the basis for communication and interaction. Images are also the natural language designers speak. The *Bridge* provides an environment that designers can naturally operate in. It supports the use of images in different modes of communication. For instance, the *meeting area* within the *Bridge* facilitates image-based



on-line meetings in a synchronous communication mode, while the *image pin-up area* is used for asynchronous idea exchanges.

The idea behind visual information exchange is to be pursued further in the development of a visual-based chat system. This synchronous communication method is to be developed to cover the necessity of on-line synchronous meetings. In such a system, the participants should be able to express their ideas and thoughts by either selecting images from a pallet of existing images or dynamically uploading their new (scanned or drawn) images from the desktop to contribute spontaneous ideas in the design discussion. All new contributions during synchronous sessions will also be added to the image library and can be used to facilitate future communications or as references.

As it is very common for people to just have an e-mail client open in normal circumstances, the *Bridge* also provides an option for its messages to reach team members via regular e-mail clients for keeping everyone updated with the most recent contributions. Users could also conveniently send email to the *Bridge* to post messages without having to go through the hassle of logging on for full system access. The key here is that everything that is exchanged through the system will be retrievable and available in the environment. The *Bridge* thus operates also as a hub to channel all types of on-line communications.



A bigger vision is for the *Bridge* to become a generic environment that supports a variety of design and art disciplines. This will inevitably involve people who have no technical background. The *Bridge* therefore, is to be engineered in the most flexible manner so as to provide non-programmers with the ability to easily create and customize instances of distance learning environment.

See also the following web pages.

- for the original beta version: <http://158.132.242.177/>
- for the web page describing the project with original *Bridge*: <http://sdalpha.polyu.edu.hk/ds4/>
- for the current development of the project: <http://158.132.242.177/sdvds/loginapp/loginform.cfm>



# The Evolution of Animation: Bedrock Revisited

Sandro M. Corsaro  
sandro corsaro animation

## 1. Introduction

Flash is the future of television animation. It cuts costs, saves time, and empowers the artist. There is a gap in the education and understanding of how to effectively utilize this program for broadcast animation purposes. This lecture will serve to educate this gap from both an artistic and production point of view. Creating broadcast animation has long been limited to having lots of money and time. The average half hour cartoon runs between 300,000 to 1.5 million dollars to produce. The traditional animation process takes over twelve weeks to produce one episode with most of the grunt work being done overseas. Flash is an artistically empowering program that will change the face of animation within the next year. Flash is slowly creeping into broadcast. Some traditionalists have hesitated to accept Flash animation as viable method of production is because the taint from the dot com era. The reason Flash has not been accepted as a viable mainstream production method is because for the most part, the animation that has been produced on the web lacked the quality of television and film. Currently, there are very few animators who have walked on both sides of the fence. But any of them will tell you Flash Animation will become a necessity in the television industry.

## 2. Overview

The presentation will clearly define the benefits using a program such as Flash MX from a production standpoint. Artistic examples such as character design, storyboarding, and animating straight into Flash MX will be covered. The bulk of the presentation will explore the business side of television animation and the archaic system that have been in place for decades. The traditional production system of animation will be starkly contrasted with the time saving efficient workflow digital animation can provide to studios.

## 3. The Evolution of Scooby Doo

In 1957, the last of the big studios closed the doors of its television animation division. The cost of a six-minute *Tom and Jerry* short was escalating at over \$50,000. MGM had no choice but to follow Disney and Warner's lead. By 1960 most animators in Hollywood were out of work. It would be two men named William Hanna and Joseph Barbera that would not only save the genre but also reinvent it. The two men invented the limited animation system. The system was based on reusing animation, strong staging, and simple gags. It would be fifty years later the same principles would become the principles of web animation. By economizing the artwork Hanna Barbera dramatically drove down the cost of production and were able to bring hundreds of artists back to their drawing boards.

It is an amazing story. What was even more amazing was my encounter with Scooby Doo creator, Iwao Takamoto six months ago. The very principles that drive the Flash animation system on the web are the ones that made Fred Flintstone a household name.

When I showed him MX he got it right away, after all it's really a system he helped pioneer.

## 4. The Future of TV Animation Production

Television animation is at a crossroad once again. Recently, an animation colleague of mine walked out of Comedy Central disappointed. His demo pilot was received well by the brass of the company but the cost to produce his style of animation for full season far exceeded what the studio was willing to pay. The next week we had a conversation. The word "Flash" was mentioned to me followed by a barrage of production and technical questions. There is an undeniable hunger to learn how to produce television animation using this program.

Almost all animated television is now done of overseas much to the dismay of many Hollywood animators. Flash MX could bring production back to the United States because it offers a cheaper and much faster way to create animation. A revolutionized system based on what Hanna Barbera introduced in the 1960s will have to emerge. This presentation would effectively show how such a system could be set up from a production point of view. There are so many animators and producers who have heard of Flash, but do not understand its capabilities both from a technological and artistic standpoint. This lecture will serve as education to creative individuals and studios to effectively drive down the cost of production. In turn, bringing jobs back to the US and saving enormous amounts of time.

## 5. About the Presenter

Specializing in pure Flash animation, Sandro Corsaro has created projects for clients such as Intel, McDonalds, MCA Records, Nestle, and E Music. Currently based in Los Angeles, he finds time for personal projects between commercial assignments and teaching Flash Animation seminars nationwide. Locally, he has taught seminars for Art Center of Pasadena and for Lynda.com's Ojai Digital Art Center. Among his latest projects is a book for New Riders Publishing entitled, *The Flash Animator*, in which he shares his techniques for applying traditional animation principles in the Flash MX environment.

Prior to his introduction to Flash, Sandro earned degrees in both Fine Arts and Animation from the University of Southern California. There he honed his traditional animation skills under the influences of great animation directors such as Tom Sito and Brad Bird. While still in college, he had the pleasure of working on the production of two animated films for Warner Bros; *The Iron Giant* and *Osmosis Jones*.



---

Contact  
info@sandrocorsaro.com  
www.sandrocorsaro.com



# The Raisio Archaeology Archive: Using Design to Build Collaborations

Lily Diaz-Kommonen  
Media Lab, University of Art and Design Helsinki  
135D Hämeentie SF 00560 Helsinki, Finland  
diaz@uiah.fi

## Introduction

Sometimes web sites get built to promote a company or a product. It is also possible that a website is constructed as part of the documentation strategy in a project. In my presentation I want to talk about how a website can be used to foster collaboration among the partners in a research project.

The web site in question is the *Raisio Archaeology Archive*. The site, which is part of my work towards the doctor's degree in design, was created with the assistance of personnel from three different institutions, namely the Media Lab of the University of Art and Design Helsinki/UIAH; the University of Turku, department of archaeology; and the city of Raisio. The web site is installed permanently in the Harkko Museum in Raisio, Finland.

## Digital archive design

Archival documents arise as a result of human activity. They constitute the records of what transpired. In this manner the Raisio Archaeology archive is evidence of the activities of a group of archaeologists and a designer working with new media technology in the transitional period between the end of the 20<sup>th</sup> and the beginning of the 21<sup>st</sup> centuries. The archive, however, does not intend to tell one particular narrative, but rather, to provide a platform of resources. Other parties, such as researchers, educators, museum personnel, and the public, can use it to tell *many narratives*.

The website was designed following models borrowed from traditional archival practices. The graphical user interface, for example, follows the form of an inventory register, a tool used by archivists when dealing with “the quantity as well as the complex organization of archival records.” (Fox and Wilkerson 1998)

The navigation strategies used are not always straightforward and different models are proposed. Illuminated letters, for example, allow the guest to navigate from a 2D document page into an interpretive 3D representation (3D World) of the landscape. This representation is based on an existing map, dating to 1725, that is also included in the archive.

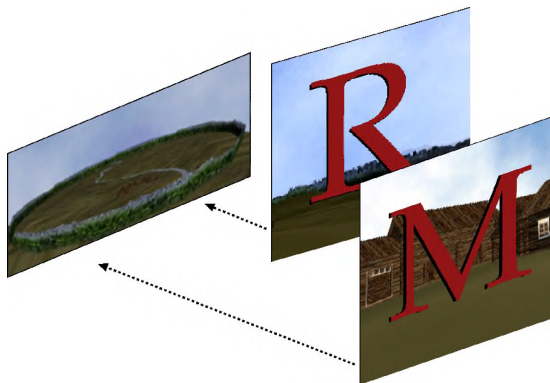


Fig. 1: Illuminated letters “R” and “M” provide entry into a 3D World.

## Communication and design processes

The resources created are basically of two types: *Media artifacts* and *tools*. The designer developed the term *media artifact* as placeholder for digital representations. It is a way to translate technical language and design processes into a more comprehensible and usable framework for the content producers. There are *media artifacts* of material culture from Southwestern Finland in the form of still images, interactive panoramas, time-based digital videos, 3D models, and textual narratives and reports. The *tools* in the archive allow the guest to augment and customize the existing interface. The *Classification* system and *Search* engine, for example, allows the guest to access a textual mode of navigation. While the expert, who knows the proper terminology can find out if a given topic is covered, the non-expert can search through an expanding tree structure and see what is indeed covered.

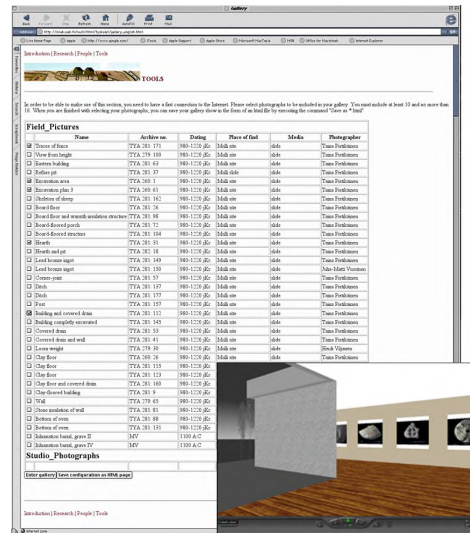


Figure 2: 3D Gallery is created dynamically. Clicking on the works displays the web page of each object in a separate window.

The 3D Gallery authoring system is a tool that allows the guest to select from the materials in the archive and curate an exhibition. The selection—that is displayed in a VRML 3D model—can be visited immediately. Or it can be saved as an HTML document for annotation and later use.

Web site Url: [http://www.mlab.uiah.fi/mulli/e\\_index.html](http://www.mlab.uiah.fi/mulli/e_index.html)

FOX MICHAEL J., & WILKERSON PETER L. 1998.  
*Introduction to Archival Organization and Description*,  
The Getty Information Institute, Suzanne Warren, Ed., 20-21.

Special thanks to Janne Pietarila.

# The Reality Cluster - Realtime Multimedia Communication with Persistence

Branden Hall  
Fig Leaf Software

Samuel Wan  
University of Michigan School of Information

## Abstract

The behavioral and cognitive principles of collaboration are well understood, i.e. how people negotiate common meaning in order to work together. During the implementation of collaborative systems, however, the significance of these principles in human-computer interaction are often shadowed by the low-level challenges of building networked applications. The REALITY CLUSTER project explores the question "What if building networked applications was easy?" by utilizing new technologies recently introduced by Macromedia and the Flash 6 plugin.

The REALITY CLUSTER allows multiple users to manipulate a graphical representation of both real-time and stored information in a common repository. The user interface for REALITY CLUSTER borrows principles found in information visualization literature to show relationships between multiple nodes of information while providing users with both focus and context in navigating the nodes. Each node may consist of either recordings or real-time channels for video, audio, text, and static graphics. Hopefully, the REALITY CLUSTER prototype will open web developers to new perspectives in designing web applications. We believe that these technologies from Macromedia, combined with strong grounding in HCI principles and software engineering, will fulfill the promise of a truly disintermediated network communication.

---

# The Xj3D Browser: Community-Based 3D Software Development

Alan D. Hudson  
Yumetech, Inc.  
giles@yumetech.com

Justin Couch  
Yumetech, Inc.  
justin@yumetech.com

Stephen N. Matsuba  
Yumetech, Inc.  
stephen@yumetech.com

## 1 Introduction

This presentation outlines the development process of the Xj3D browser. Xj3D is an open source API for developing X3D and VRML 97 applications. It is also the sample implementation and test bed for the next generation VRML specification known as Extensible 3D (X3D). Indeed, Xj3D was initiated by the Web3D Consortium to provide input to the X3D authors and the 3D graphics community with input concerning problems and ambiguities with the specification.

## 2 The X3D Specification

X3D is an open, royalty-free standard for networked, real-time 3D graphics. People in the VRML community were given the opportunity to provide input through the Web3D Consortium's various mailing lists. Moreover, companies established their own working group to identify the features they required in the specification. As the process developed, however, it became apparent that a small team of contributors would form the core participants.

From the beginning, it was clear that X3D involved a major reconsideration of the VRML97 specification. Various parties expressed competing desires. Initially, work on X3D began with a focused group of people selected from existing VRML-based companies. However, this effort collapsed due to economic problems. After an extended period of slow progress, a collation of academics, smaller companies and individuals was formed to move the specification process forward.

## 3 The Goals of the Xj3D Browser

Xj3D was started with a donation of source from Sun Microsystems in 1998 and currently is a project of the Web3D consortium's source task group. The code-base has four main goals. The first is to discover and identify problem areas in the X3D specification. The second is to provide a platform for the community to test new ideas for possible incorporation into future specifications. The third is to make it easier for other software to import and export 3d graphics data. The final goal is to develop a production quality browser.

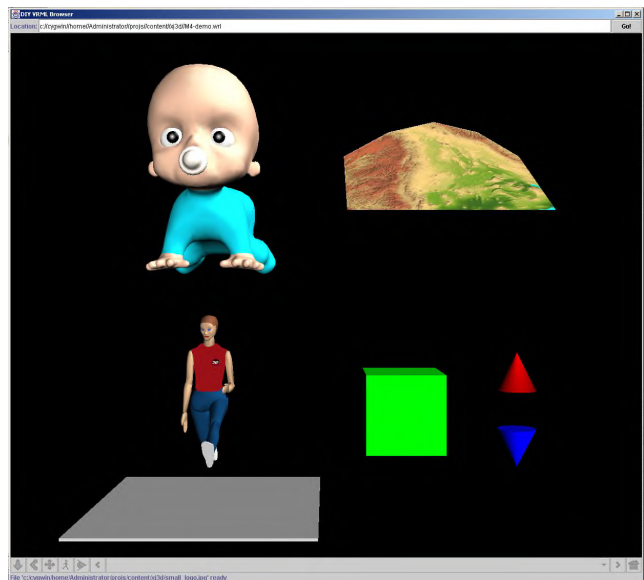
## 4 The Xj3D Development Process

Originally, Xj3D was based on source code for a VRML 97 browser that was developed at Sun Microsystems. The Web3D Consortium designated the Source Code Management Working Group headed by Alan D. Hudson of Yumetech, Inc. as the body responsible for developing a code base that would become the test bed for the X3D specification.

The Web3D community believed strongly that the development of Xj3D had to be an open process that was linked to the evolution of the X3D specification. To achieve this goal, the development of the code-base would be managed over the Internet. The Source Management Working Group established an online CVS repository where individuals working on the code would be able to send and track any changes. A mailing list was

opened to allow interested parties to participate in creating Xj3D, and the web3d.org website made all the latest releases available for download..

Early in the process, however, it became apparent that the project required more impetus from a central source. Yumetech took on the role of lead developer and project supervisor. The company's lead programmers performed an assessment of the original code-base and determined that a completely new base for Xj3D would be necessary to make it easier to use and more efficient. The M2 release was completed for the 2001 SIGGRAPH Conference. Since then, Yumetech has overseen and participated in developing two more releases.



## 5 Assessment of the Community-Based Development Model

The community-based development model for the Xj3D source code has proven to be successful. The Source Code Management Working Group list and the Java 3D mailing lists provide a wealth of suggestions, feature requests, bug reports and fixes that would not be available if development were done in a closed environment

Moreover, the open process used to create Xj3D has directly influenced the evolution of the X3D standard. As different elements of the specification are implemented in the browser, ambiguities and problems in the specification become apparent. The Xj3D browser developers are able to point to specific parts of X3D and offer changes to improve it. Thus the Xj3D development process provided a real grounding for the development of the X3D specification.

## References

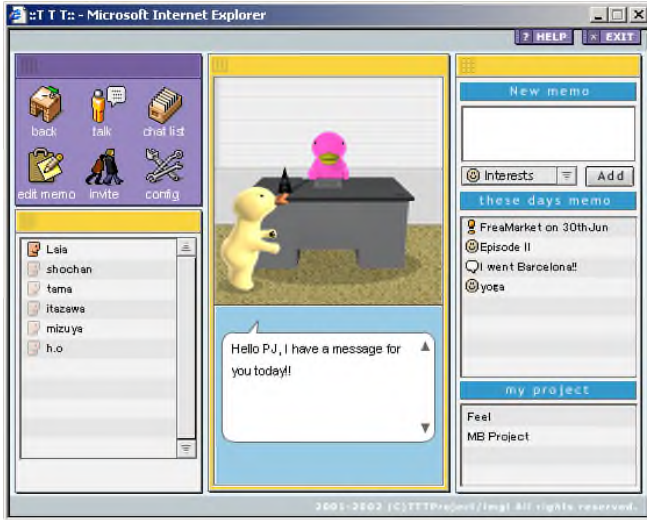
"Xj3D Open Source VRML/X3D Toolkit."  
[http://www.web3d.org/fs\\_workinggroups.htm](http://www.web3d.org/fs_workinggroups.htm).

# TTT:A Web community tool mediated by friends

Yuichiro HARAGUCHI  
Keio University  
hrgci@imgl.sfc.keio.ac.jp

Sakura TOYABE  
Keio University  
saqura@sfc.keio.ac.jp

Masaru MURATA  
Keio University  
m\_m@imgl.sfc.keio.ac.jp



## I. Introduction

TTT is a Web community tool that promotes new encounters mediated by friends.

Participating in a network community like BBS or a mailing list, you can find people who share the same interests as yourself. However, on the Internet, it is difficult to make friends with whom you can trust because conventional Web community tools lack the structures for creating trust between its members; an important required element of community building in the real world. Like a party or an exchange meeting, TTT makes it possible to extend human connections by utilizing what we believe is the source of "trust", that is, friendship. This paper illustrates how this system is designed.

## II. Communication Design for TTT

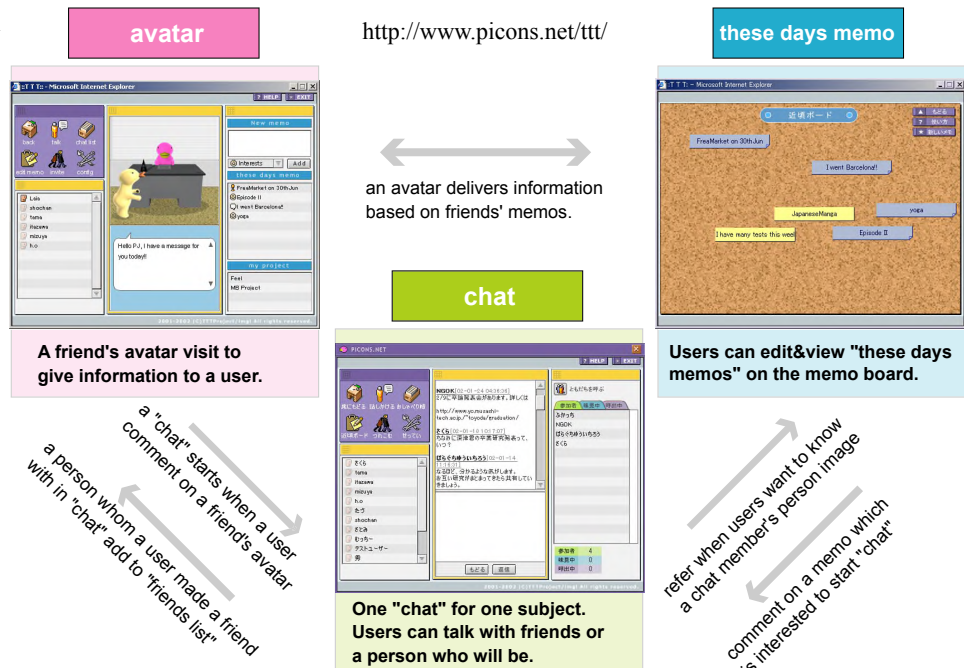
We modeled the processes of "trust building" and "encountering" based on everyday experience. The model was constructed from the following cases.

### 1. Letting an acquaintance know personal interests or news

-- My friend A and I often talk about recent news and mutual interests.

### 2. Information transmission via friends

-- One day I told A that I am recently interested in X. Then A told his friend B who has a similar interest that I am interested in X.



### 3. Encounters with people of similar interests and values

-- Through A, B and I had a chance to meet and talk about our interests. Consequently, we became friends and often talk about X and other personal news.

In the TTT system, personal news information called "recent memo" is shared among friends, and a community unit called a "chat" is built when someone comments on it. In addition, "recent memo" can be passed on to a friend's friend when mediated by a friend. Hence, the burden of trust creation is reduced, and we are able to meet people who have similar interests and concerns. Moreover, since the system is based on friendship in the real world, it is possible to actually meet and to talk to new friends offline. A community on TTT is certain to be a powerful because it unites the real world and the virtual world.

## III. What are the differences from conventional communication tools?

- Using friendship information, we can achieve both "trust" and "new encounters".
- It is easy for users to start a conversation because the only requirement for participation is sharing personal news with friends.
- A unique avatar delivers information instead of users.

## IV. Technologies

In order to perform real-time chat and add animation to our avatar within a web browser, we used Macromedia's ActionScript(client side) and PHP(server side).

## VI. URL

<http://www.picons.net/ttt/>



# VisSheet Redux: Redesigning a Visualization Exploration Spreadsheet for the Web

T.J. Jankun-Kelly and Kwan-Liu Ma  
University of California, Davis  
{kelly,ma}@cs.ucdavis.edu

## Introduction

The exploration of complex data sets requires interfaces to present and navigate through the visualization of the data. In recent work [Jankun-Kelly and Ma 2001], we produced a visualization exploration spreadsheet to address this issue. The developed application, however, was implemented for off-line use only. For data sets on remote sites, this approach is not appropriate. Thus, a web-based version of the visualization exploration spreadsheet is needed. This abstract discusses the process of transforming the interface from an off-line to an on-line design.

## From Off-line to On-line

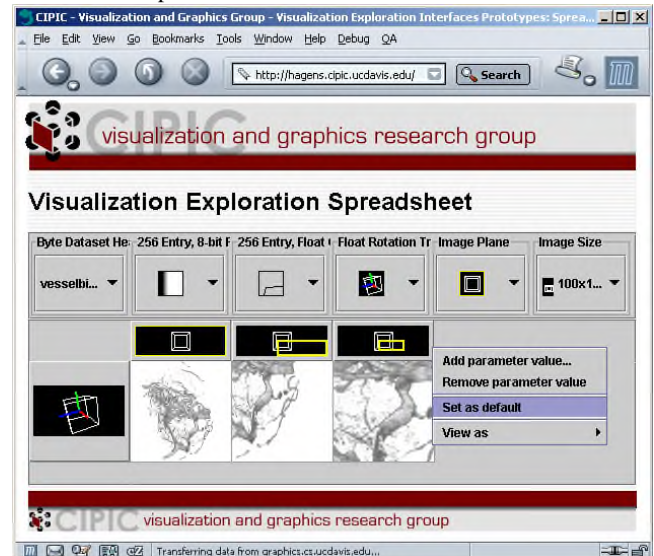
Our spreadsheet-like interface (VisSheet, for short) was designed to assist visualization exploration by providing context for where the user is in their exploration, where they have been, and suggesting where they may go next. The VisSheet addresses these tasks by providing a movable, scalable window into the visualization parameter space. By manipulating the visualization parameters, the user changes the position and size of this window. Only two visualization parameters are displayed at a time: one along the rows and another along the columns. For the non-displayed parameters, a set of default values is maintained which may be updated at run-time. Parameter values are rendered as glyphs. Cells—representing a combination of the row, column, and default parameter values—display the visualization results. By changing the the default values for non-displayed parameters or which parameters are displayed along the rows or columns, the “window” can be moved in the visualization space. Thus, the data exploration process becomes the process of manipulating the spreadsheet window through the visualization space.

The original design of the VisSheet consisted of three major components: the active View on the visualization parameter space, the Model of the visualization transform responsible for the parameter space, and a Session which records the user’s process. These components were all designed to be modular. For example, one could change the type of visualization performed by replacing the visualization transform encapsulated by the Model object. This module handles the rendering requests. By replacing the Model object with a version that performed remote visualization, the application could become partially on-line. This configuration, however, was fully “web-aware”: it could neither be accessed on-line nor visualize data from different servers. Thus, a redesign was performed.

There were two major goals for the redesign effort. The first was to allow the VisSheet to be “fully” on-line. By this, we mean that the application should be accessible over the Internet on a wide variety of machines. The second goal was to maintain modularity. The VisSheet is only one of the exploration interfaces being developed. It is important that the reusable components common to the other interfaces—such as the modules capturing the exploration process—could be reused. The modularity also extends to the visualization itself: Like the original VisSheet, the new version must be able to interface with different visualization transforms.

The first goal determined a series of constraints on the new design. Since the application was to be on the web, it was assumed that there would be no interaction with the local machine. In addition, it is desirable that users of the on-line VisSheet have a very low barrier of entry. Fewer constraints were due to the second goal. At the time of the VisSheet redesign, the modules forming the core of the new visualization exploration framework were also be written. Thus there was no legacy code to support in the new implementation.

The final design of the system consists of rewrites of the original three major systems. The visualization framework itself has been implemented in Python. In the new framework, an abstract, UI-toolkit independent VisSheet module has been developed; for the web VisSheet, a Java interface is provided by using Jython, a Python environment for Java. In our prototype, the VisSheet applet, written in Java and Jython, communicates to a volume visualization server written in Python with a C++ hardware accelerated renderer. This interoperation of languages attests to the modularity of the system. This system will be demonstrated, detailing both how the framework assists remote visualization and how the system evolved to this point.



Oliver Kreylos provided the texture-based volume renderer for our use.

## References

- JANKUN-KELLY, T. J., AND MA, K.-L. 2001. Visualization exploration and encapsulation via a spreadsheet-like interface. *IEEE Transactions on Visualization and Computer Graphics* 7, 3 (July/Sept.), 275–287.

# Wegzeit - the Geometry of Relative Distance

Dietmar Offenhuber

TU Vienna / Ars Electronica Center \*

## Abstract

This web3d project explores how the concept of non-linear space - that is space structured by relative units - can be used in VR and architecture. It offers a dynamic view on Los Angeles' structure, radically different from usual architectural representations.

## 1 Introduction

The project presented here explores how non-linear space-space that is structured by relative units-can be used in VR and architecture. It offers a dynamic view of Los Angeles' structure that is radically different from conventional architectural representations. The background of the project is described in more detail in the sketches&application presentation of this project.

## 2 Implementation

The project consists of six dynamic virtual environments that propose models of how to visualize three-dimensional relative spaces. They deal with certain properties and effects caused by the nature of relative space such as the asymmetry of temporal distances.

The environments were designed in such a way that they can be used both in immersive VR and web3d contexts. All deformation models are managed by scripted behaviors that are evaluated in real-time. Most of the underlying parameters can be changed interactively.

The creation of the environments involved little modeling; space is constructed by the actual visual perception of the driver. The frontages of the streets examined are represented by long strips of texture that were created from videos taken from the moving car. Three of these models can be briefly described here:

**Linear distance model** Space is described here as a path-network with variable distances between its nodes. A number of constraints give more control over the resulting shape. In the example, a half-mile area is displayed in time-space based on speed limits, phases of traffic lights and actual driving time.

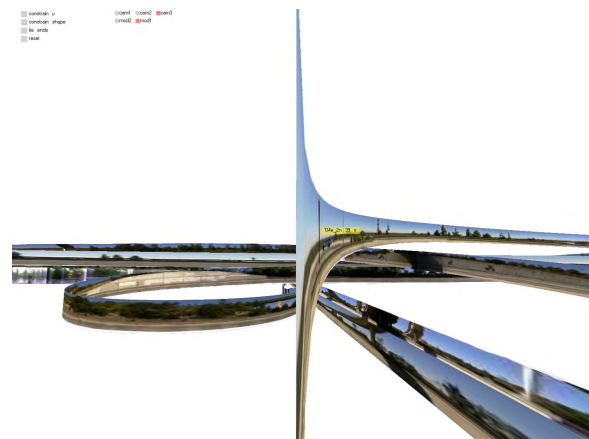
**Area model** While the network model works fine for the representation of relative distances, other relative parameters are based

on surface units. For these parameters, a different model is necessary. The example presented investigates the space of attention. The diagram of all installed written signs on the Las Vegas Strip from the essay "Learning from Las Vegas" by Robert Venturi et al. is transformed into relative space based on local information densities. Areas with high information density seem inflated.

**Slope model** Temporal distances are usually asymmetric; the distance from point A to point B might be considerably different than the distance from B to A. This asymmetry can be expressed as steepness of the terrain. If travelling in the opposite direction is much faster, the movement is somehow against the stream-like going up a steep mountain. If we are on the faster track and it's more difficult to go in the other direction, it's like going downhill. This metaphor is the basis for the following environments. The model is influenced by M. Fujihata's Mt. Fuji Project, which uses a similar approach to reconstructing the shape of Mt. Fuji from the walking-velocity of a group of people.

One of the biggest challenges in working with non-linear space in VR is maintaining its legibility for the user. Working carefully with internal and external views of the environment, constraining the user's motion to a set of paths, and simplifying interaction are possibilities to prevent confusion in an environment without any kind of static form.

The project can be viewed at <http://futurelab.aec.at/wegzeit>



\*e-mail:didi@aec.at

