Mitsubishi Electric Research Laboratories (MERL)

Annual Report

July 1999 through June 2000

TR2000-00

Welcome to MERL – Mitsubishi Electric Research Laboratories, the North American corporate R&D arm of Mitsubishi Electric Corporation (MELCO). In this report you will find descriptions of MERL as a whole, and of our three laboratories, MERL Cambridge Research, MERL Cambridge Systems, and MERL Murray Hill . We also describe MERL Concord, which is in the process of moving into a separate business unit.

201 Broadway, Cambridge, Massachusetts 02139
617.621.7500

Production:
Karen Dickie, Janet O'Halloran, Richard Waters

Letter from the President

Mitsubishi Electric Research Laboratories (MERL) is the Corporate R&D arm of the Mitsubishi Electric Company (MELCO) in North America.

MERL conducts application-motivated basic research and advanced development in computer and communications technology. We seek to contribute to the advancement of science and to meet human needs by creating fundamental new technology that expands the productive use of computers. Our vision is both long-term and market-driven. We seek to anticipate and then meet market needs and business opportunities. This means exploring entirely new possibilities as well as improving what is now possible. Because computer and communication technology affects nearly every aspect of modern life, and because MELCO is a broadly diversified company, there are many opportunities for our results to contribute to MELCO's business success.

MERL focuses on three key technology sectors: human/computer interaction (HCI) featuring advanced graphics and computer vision technologies, Internet computing and applications, and digital communication. The labs within MERL focus on specific segments of these technologies, while working collaboratively to achieve groundbreaking results. Our output ranges from papers and patents, through proof-of-concept hardware and software prototypes, to industry-first products.

MERL is small enough to be agile and flexible in the dynamic marketplace of ideas and innovations. In addition, we gain leverage from the size, recognition and diversity of our strong global parent. We turn our technical achievements into business successes by partnering with MELCO's business units and with other labs in MELCO's global R&D network. With selected projects, we do early business development directly with customers, in collaboration with MELCO and/or other partners.
Located in North America, we perform our R&D in the world's liveliest hotbed of computer technology innovation and business opportunity. We interact with North American MELCO subsidiaries and customers. We are strongly involved in THE R&D community and standards activities, maintaining long-standing cooperative relationships with research universities such as Brown, CMU, Georgia Tech, MIT, the University of Michigan, Princeton, SUNY-Stony Brook, and the University of Massachusetts. We encourage our staff to be involved in their professional communities via conferences, papers, reviewing, and continuing professional development.

Please visit us in person or via our web site "http://www.merl.com".

Richard C. Waters
President and CEO

Table of Contents

Mitsubishi Electric Research Laboratories

With its Headquarters in Cambridge MA, Mitsubishi Electric Research Laboratories (MERL) consists of four laboratories, two in Cambridge, one in New Jersey and one outside of Boston, MA.

- MERL Concord is a hardware and software development laboratory. MERL Concord designed the first ASIC chip capable of real-time volume rendering on a PC plug-in board. The medical and geophysical industries are targeted as early users of volume graphics. A second-generation chip is under development for shipment next year. MERL Concord provides advanced development and engineering support to Real Time Visualization, a new MELCO business entity responsible for volume graphics.
- MERL Cambridge Research is a basic computer science research laboratory, with diverse research specialties ranging from real-time distributed systems through computer graphics, and computer vision to human/computer interaction. MERL Cambridge Research's Collagen system for collaborative interface agents is being used by a number of research institutions and is being evaluated for use in MELCO products. The lab's work on digital audio has being incorporated into the evolving MPEG-7 standard. Graphics research at the lab created and continues to support MERL Concord's product development activities.
- MERL Cambridge Systems is an advanced development laboratory specializing in Internet software and home automation. MERL Cambridge Systems' Concordia system supports mobile software agents in Java. Concordia is being used by several groups within MELCO and has been licensed to two companies in North America. A key future direction for MERL Cambridge Systems is work on home automation applications in conjunction with the other MERL labs.
- MERL Murray Hill is an advanced development laboratory specializing in digital TV, digital communication and the convergence of television, communications, and computer technologies. MERL Murray Hill works closely with MELCO's semiconductor, communications, and living environment & digital media equipment business units. In partnership with Lucent and MELCO laboratories in Japan, MERL Murray Hill designed the world's first chip set capable of receiving US HDTV boradcasts.

Within MERL, we emphasize collaboration among the four labs. The integration of the expertise of multiple labs and a partnership between research and advanced development provides a powerful basis for industry first developments. Our work on volume graphics and our beginning work on home automation are prime examples of how this kind of collaboration can produce successful innovation.

Management Team

Mitsubishi Electric Research Laboratories
President & CEO: Dr. Richard (Dick) Waters
EVP, CFO & CLO: Mr. Takashi (Tak) Hiratsuka
EVP: Dr. Tommy Poon

Concord Ms. Beverly Schultz	Hardware Development Software Development
Cambridge Research Dr. Richard (Dick) Waters	Computer Graphics & Vision Human/Computer Interaction
Cambridge Systems Dr. Douglas Kleinmann	Mobile Agent Technology Home Automation
Murray Hill Dr. Tommy Poon	Digital Communications Digital Video
Marketing & Business Development Mr. Adam Bogue	
Patents Mr. Dirk Brinkman, Esq.	
Accounting Ms. Kathleen McCarthy	
Administration Ms. Janet O'Halloran	
System Administration Mr. Erik Piip	

Richard (Dick) Waters

President & Chief Executive Officer, Research Fellow & Director Cambridge Research
617-621-7508, waters@merl.com

Dick Waters received his PhD in artificial intelligence (AI) from MIT in 1978. For the next 13 years he continued to work at the MIT AI Lab as a Research Scientist and co-principal investigator of the Programmer's Apprentice project. Dick was a founding member of MERL Cambridge Research in 1991. As a research scientist, his work centered on multi-user interactive environments for work, learning and play. In January 1998, Dick became director of MERL Cambridge Research. In December 1999, he became CEO of MERL as a whole as well.

Takashi (Tak) Hiratsuka
Executive Vice President, Chief Financial Officer & Chief Liaison Officer
617-621-7513, hiratsuka@merl.com

After many years in MELCO's computer business unit, Tak Hiratsuka came to the US in 1988, as the liaison between MELCO and MELCO's subsidiary Horizon Research Inc. After 6 years there, Tak took on other duties for MELCO including a stint as Assistant Deputy Director of Corporate R&D in Tokyo. In December 1999, Tak returned to the US as MELCO's Liaison to MERL. Tak plays an essential role in building partnerships between MERL and MELCO.

Tommy C. Poon
Executive Vice President; Director of Murray Hill
908-665-1201, poon@merl.com

Tommy Poon received his PhD from Columbia University in 1980, and worked briefly at RCA Laboratory before moving to AT&T Bell Laboratories. During 13 years with Bell Labs, Tommy managed numerous R&D projects in telecommunications and signal processing. He joined MERL Murray Hill as Director in 1995. His primary research interests include digital and wireless communications, digital video and digital networks.

Douglas E. Kleinmann
Vice President; Director Cambridge Systems
781-466-8322, kleinmann@merl.com

Doug Kleinmann received his PhD in Astronomy from Rice in 1968. He spent 11 years as an astronomer, principally at the Smithsonian Astrophysical Observatory at Harvard. In 1980 he received an MBA from the MIT Sloan School and began an 18-year career as a Program Manager at Honeywell/Loral/Lockheed Martin. In 1999, Doug received an MS in software engineering from Brandeis. He joined MERL Cambridge Systems as director in April, 2000.

Beverly J. Schultz
Vice President of Engineering Concord
978-369-6500, schultz@merl.com

Beverly Schultz holds an MS in Computer Science from the University of Dayton. Beverly's 20+ years of experience include being VP of Software Engineering for Avid Technology, building video editing systems and VP of Engineering for Number Nine Visual Technology, moving the company from 2D to 3D graphics. In early 1998, Beverly joined MERL Concord as VP of Engineering.

Adam Bogue
Vice President of Marketing and Business Development
617-621-7590, bogue@merl.com

Adam Bogue received his BS in Materials Science and Engineering from MIT in 1986 and an MBA from the MIT Sloan School in 1990. For 3 years at GenRad inc, he was responsible for managing a new line of automatic test equipment. Adam spent the last 7 years at Active Control eXperts Inc. beginning as Director of Sales and Marketing and ending as Vice President, Core and New Business Unit. Adam began work for MERL in June of 2000.

Mitsubishi Electric

Number 92 on Fortune magazine's list of the world's 500 largest corporations, Mitsubishi Electric Corporation (MELCO) has approximately $35 billion in annual sales and 140,000 employees in 34 countries. Like most Japanese companies, the lingering malaise of the Japanese economy has affected MELCO. However, MELCO returned to profitability in 1999 after losses in 1997 and 1998.

MELCO is organized into ten business units. Because information technology is important to each of these business units, MERL works with them all. (The rightmost column below shows the abbreviated Japanese business unit names commonly used by MELCO insiders.)

Mitsubishi Electric Diversified Electrical and Electronics Manufacturer		"MELCO"
Information Network Services Corporate Intranets, Internet Service Provider	**(Tokyo)**	"Inethon"
Social Infrastructure Systems Power Equipment, Plant Control, Transportation	**(Kobe, Itami, Nagasaki)**	"Shakaihon"
Building Systems Elevators, Escalators, Building Monitoring	**(Inazawa)**	"Biruhon"
Electronic Systems Satellites, Radar, Military Systems	**(Kamakura, Itami)**	"Denshihon"
Communication Systems Wired Communications, Broadcast Communications, Cell Phones	**(Kamakura, Itami)**	"Tsuhon"
Information Systems Turnkey Information Systems, Computer Hardware	**(Ofuna)**	"Johon"
Living Environment & Digital Media Equip Air Conditioners, Refrigerators, TVs, DVDs, LCD Projectors	**(Shizuoka, Kyoto)**	"Lihon"
Factory Automation Programmable Logic Controllers, Industrial Machine Tools	**(Nagoya)**	"FAhon"
Automotive Equipment Alternators, Engine Controllers, Car Stereos, Car Navigation	**(Himeji, Sanda)**	"Shahon"
Semiconductors Memory, Microcontrollers, Optical Semiconductors, LCD Panels	**(Kita-Itami, Sagamihara)**	"Hanpon"

It is worthy of note that there are over 30 major independent companies that use the word "Mitsubishi" in their names and use the red three-diamond logo, including Mitsubishi Motors, Mitsubishi Heavy Industries, Bank of Tokyo-Mitsubishi, Mitsubishi Chemicals, and the Mitsubishi Trading company. These companies have shared roots in 19th century Japan; however, these companies have been separate from each other for many years and MELCO has been separate from all of them since its founding in 1925.

Many of MELCO's business units have North American subsidiaries. These subsidiaries, particularly those that do design as well as manufacturing and sales, are natural partners of MERL.

Mitsubishi Electric United States, Inc. (MEUS)
Sales: Several BUs (Los Angeles, Sunnyvale & other cities)
Semiconductors, Air Conditioning, Elevators

Mitsubishi Wireless Communications, Inc. (MWCI)
Design, Manufacturing & Sales: Communications BU (Atlanta, San Diego)
Cell Phones

Mitsubishi Digital Electronics America, Inc. (MDEA)
Design, Manufacturing & Sales: Living Environment BU (Los Angeles, Mexicali MX)
High Definition Projection Televisions, DVDs, VCRs

Mitsubishi Electric Air Tech America (MELAIR)
Manufacturing & Sales: Living Environment BU (Madison WI)
Energy Recovery Ventilation Systems

Mitsubishi Electric Automation, Inc. (MEAU)
Sales & Installation: Factory Automation BU (Chicago)
Factory Automation Equipment

Mitsubishi Electric Automotive America, Inc. (MEAA)
Manufacturing & Sales, Automotive Equipment BU (Detroit, Mason OH)
Auto Parts

Mitsubishi Electric Power Products, Inc. (MEPPI)
Design, Manufacturing & Sales: Social Infrastructure BU (Pittsburgh)
Power Transmission Products

NEC-Mitsubishi Electronics Display of America, Inc. (NMEDA)
Manufacturing & Sales: 50% owned by MELCO (Itasca Il, Mexicali MX)
CRT and LCD Computer Monitors

Mitsubishi Electric Corporate R&D

MERL is one of six laboratories in MELCO's Global Corporate R&D network. The chart below summarizes the primary activities of these labs. MERL pursues collaborations with all these labs. (The rightmost column below shows the abbreviated Japanese laboratory names commonly used by MELCO insiders.)

Corporate R&D Headquarters		"Hatsuhon"
Managing MELCO's R&D:	25 people (Tokyo)	
Managing Director:	Dr. Tamotsu Nomakuchi	
General Manager:	Dr. Ken Satoh	

Advanced Technology R&D Center (ATC)		"Sentansoken"
Research & Advanced Development:	500 people (Itami, near Osaka)	
General Manager:	Dr. Goro Yamanaka	
Materials, Semiconductor Devices, Electrical & Mechanical Engineering		

Industrial Electronics & Systems Laboratory (IESL)		"Sanken"
Advanced Development:	250 people (Itami, near Osaka)	
General Manager:	Dr. Tadatoshi Yamada	
Power Electronics, Plant Control, Factory Automation, Building Systems		

Information Technology R&D Center (ITC)		"Johosoken"
Advanced Development:	800 people (Ofuna, near Tokyo)	
General Manager:	Dr. Hitoshi Ogata	
Information Systems, Communications, Opto-Electronics		

Industrial Design Center (IDC)		"IDken"
Advanced Development:	90 people (Ofuna, near Tokyo)	
General Manager:	Mr. Seiji Wada	
Industrial Design, Usability Studies		

Mitsubishi Electric Research Laboratories (MERL)	
Research & Advanced Development:	115 people (four U.S. locations)
President:	Dr. Richard Waters
Computer Graphics, Internet Software, Digital Audio & Video Communications	

Mitsubishi Electric Information Technology Centre Europe (ITE)	
Advanced Development:	60 people (Rennes, France & London, England)
President:	Mr. Shuichi Ono
Wireless Communications, Digital Audio & Video Broadcasting	

Historical Note

The history of MERL is a story of separate labs with separate histories coming together into an organization that is growing to become much more than the sum of its parts.

In 1984, MELCO's computer business unit founded a laboratory in Waltham MA called Horizon Research inc. (HRI). HRI was initially created to design IBM compatible computers, which MELCO then produced and sold. However, in the late 80's MELCO decided to withdraw from this business. In the ensuing years, Horizon transformed itself into a software-oriented laboratory working with multiple parts of MELCO.

In 1991, MELCO's central R&D organization founded a research laboratory in Cambridge MA called Mitsubishi Electric Research Laboratories (MERL). The charter of this lab was to do fundamental research in computer science, with the goal of leading MELCO into the future with innovative new technologies.

In 1993, MELCO's audio/visual business unit founded a laboratory in Princeton New Jersey called the Advanced Television Laboratory (ATL). ATL was initially created to develop a chip set capable of decoding US HDTV signals and for several years this was the sole focus of the lab. However, starting in the late 1990s ATL branched out into other digital video technologies and digital communications in general. To interact more easily with Bell Laboratories, ATL moved from Princeton to Murray Hill. In addition, ATL created a satellite operation in Sunnyvale CA focussing on digital broadcasting and home networking.

In 1995, HRI and ATL were transferred from their original parent business units to central R&D and central R&D reorganized its US operations. Central R&D created a new US entity called Mitsubishi Electric Information Technology Center America (MEITCA or as we more simply referred to it internally ITA). HRI was incorporated into ITA becoming the Horizon Systems Lab (ITA/HSL). ATL was incorporated into ITA becoming ITA/ATL. The following year, MERL was incorporated into ITA as well, becoming ITA/MERL.

The ITA that resulted in 1996 had a great deal of promise, but also some significant obstacles to be overcome. By combining advanced development and basic research into a single organization, ITA made it easier for technological advances to be perfected and transferred into MELCO business. Further, by combining expertise from a wide range of areas and experience with a wide range of MELCO business units into a single organization, ITA was in a good position to understand a wide range of technical trends, market needs, and MELCO business opportunities.

One of the first fruits of the synergies in ITA was the creation within ITA of the Volume Graphics Organization (ITA/VGO). The charter of this group was the design of a real time volume-graphics rendering chip that could be hosted on a PC plug in board. VGO was based on research carried out at ITA/MERL and began as a project within ITA/HSL before growing so large that it was made a separate laboratory within ITA in 1998. From its inception through 1998, ITA/VGO shared space with ITA/MERL and interacted closely with ITA/MERL researchers. However, in 1999 ITA/VGO moved to separate quarters in Concord MA as sales of its first product were beginning.

The obstacles faced by ITA were the same as those faced by any organization created by the combination of preexisting parts. The natural momentum was for ITA/HSL, ITA/ATL, and ITA/MERL to continue operating on their own following their own goals and agendas, with collaboration being the exception rather than the rule. Much time and effort was required to overcome this momentum and begin a move toward a unified organization that could achieve ITA's full potential; however, in 1999-2000 the tide began to change and motion toward a unified organization became clearly evident.

In mid-1999, ITA's staff of some 115 people was spread across four labs operating in five locations—ITA/HSL in Waltham MA, ITA/ATL in NJ and CA, ITA/MERL in Cambridge MA, and ITA/VGO in Concord MA—making collaboration difficult. Since then, the ITA/ATL people in Sunnyvale have joined the rest of ITA/ATL in New Jersey and half the people at ITA/HSL have joined ITA/MERL at the Cambridge facility. The rest ITA/HSL will make this move when more space in Cambridge becomes available. Lastly, ITA/VGO is in the process of leaving ITA and joining the Real Time Visualization Business organization that was created to market the products created by ITA/VGO. However, by mid-2001, ITA will have been transformed into a more coherent organization consisting of three labs operating in two locations (Cambridge MA and Murray Hill NJ).

A smaller but still significant obstacle to a unified ITA was its name. Unfortunately, "Mitsubishi Electric Information Technology Center America" is too much of a mouthful to be an effective name. Its acronym MEITCA is equally problematical, which lead us to use the shortened form ITA instead. However, ITA has a weak connection to the full name and is already in use by other companies, one of which uses the web address "http://www.ita.com". As a result, we were never able to make the name ITA a unifying banner for the organization. Instead, the labs in ITA continued to operate primarily under their individual names.

In the summer of 2000, we instituted a new naming strategy modeled on the approach of Bell Labs. We adopted MERL as a shorter and more pronounceable name for the company as a whole. Then, to emphasis MERL as a unified name, we switched to location-based names for the individual labs—MERL Concord, MERL Murray Hill, MERL Cambridge Research, and MERL Cambridge Systems, the former ITA/VGO, ITA/ATL, ITA/MERL, and ITA/HSL, respectively.

MERL Concord

MERL Concord was created to develop leading-edge volume graphics technology and to drive it into initial product in Q2 99. An engineering team was formed to build products that can enable compelling visual experiences of unrivaled fidelity by exploiting discrete and geometric representations. The first product is a graphics board that does volume visualization in real time. The VolumePro500 (VG500) product includes a proprietary ASIC chip that enables real time volume visualization on a board with 160Mbytes of memory, and a CD that contains drivers that work with the hardware and a Volume Library Interface to take advantage of the features of the hardware. The first product has been shipping to customers since May 99. Follow-up releases, V1.1 and V1.2 have occurred since the initial shipment, and V2.0 was released in June 00.

The first products from MERL Concord are creating and establishing a leadership position in this new field of volume graphics techniques and methods. The team is nearing completion of the design of a second chip, code named Condor, which will be sold as VolumePro1000 (VG1000). This chip will give the next level of performance and quality to the visualization of volumes. Over time, with a series of releases of software for Condor, the volume rendering technology will be integrated with traditional 3D polygon graphics, and enhancements will continue to be made to meet the needs of customers and to advance the technology offered.

MERL Concord is located in Concord, MA, and has a staff of 25. Reflecting its product development nature, this organization will soon be leaving MERL to become part of MELCO's Semiconductor business unit.

The following awards have been received from the VG500 product:

1999 Innovation Award, Computer Graphics World, January 2000.

EDN Innovation of the Year Award, Finalist for 1999.

Top 100 Hot Products for 1999, EDN Magazine.

Top 10 Products of 1999, December 1999 Nikkai Kogyo Shimbun, Tokyo, Japan.

EE Times Product of the Week, July 5, 1999.

Recent Major Publications

Bhalerao, A. Pfister, Hanspeter., Halle, M. and Kikinis, R. "Fast Re-Rendering of Volume and Surface Graphics by Depth, Color, and Opacity Buffering,", Journal of Medical Image Analysis, *To appear in May 2000.*

Pfister, Hanspeter, "VolumePro - At the Frontier of Advanced Volume Graphics", Nikkei Science, October 1999

Gasparakis, C."Multi-resolution Multi-field Ray Tracing: A Mathematical overview", Proceedings IEEE Visualization 1999, p. 199-206, October 24, 1999.

Pfister, Hanspeter, Hardenbergh, Jan, Knittel, Jim, Lauer, Hugh and Seiler, Larry "The VolumePro Real-Time Ray-Casting System" *Proceedings of SIGGRAPH* 99', p.251-260, Los Angeles, CA, August 1999.

Ray, H, , Pfister, Hanspeter, Silver, D., Cook, T. A. "Ray-Casting Architectures for Volume Visualization", IEEE Transactions on Visualization and Computer Graphics,p.210-223, vol.5, No. 3, July/September 1999.

Pfister, Hanspeter, Keynote Address at the 5th Japanese Visualization Conference, "Real Time Volume Visualization with VolumePro" - Tokyo, Japan, 1999

Pfister, Hanspeter, "Architectures for Real-Time Volume Rendering, Journal of Future Generation Computer Systems (FGCS)", *Elsevier Science* , pp. 1 - 9, Vol. 15, No.1, February 1999.

Ogata, M., Ohkami, Takahidi, Lauer, Hugh, Pfister, Hanspeter, "A Real-Time Volume Rendering Architecture Using an Adaptive Resampling Scheme for Parallel and Perspective Projections", *ACM / IEEE Symposium on Volume Visualization*, pp. 31-38, Durham, NC, October 1998.

Articles

Venturing out Mitsubishi Tests the Entrepreneurial Water, Presented in Technology Review November/December 1999.

VolumePro 1000 Expands 3D Vision, Printed in Microprocessor Report October 25, 1999.

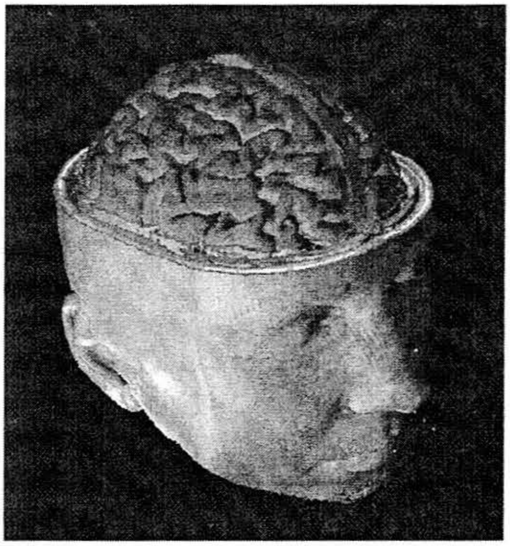
Volume-rendering PCI Board Casts Rays in Real-time,_Presented in EE Times July 5, 1999.

Mitsubishi Soups up Graphics Board to Bring 3-D Volume Rendering to PCs, Presented in Diagnostic Imaging Scan June 9, 1999.

Workstation 3-D Graphics Get a Reality Boost,Presented in EETIME March 1999 by Anthony Cataldo.

Project Descriptions

VolumePro500



With first silicon in April of 1999, the VolumPro500 was the world's first real time volume rendering support for PC's and remains the only hardware of its kind. It can render 256^3 volume data at 30 frames per second. This is much faster than any software implementation and much less expensive than any other hardware implementation.

Background and objectives: Volume rendering has matured as a method for visualizing 3D sampled, simulated, and synthetic datasets. Unlike surface-based representations, volumetric representations can embody interior structures and composition. This allows the visualization of internal structure of the 3D data, including amorphous and semi-transparent features. Additionally, the emerging field of volume graphics is using volume data for physics-based modeling of complex objects, including geometrical models that have been synthesized can be performed relatively easily with a volumetric representation. More recently, volume rendering has become essential for directly viewing changes of dynamically sampled data, for example, for the visualization of a beating heart under real-time 3D ultrasound.

Technical discussion: The genisis of the VolumePro500 began with the thesis work of Hanspeter Pfister at SUNY on technology for rendering volumes. The work then continued at MERL and matured into a development project, first at HSL and then in an organization (VGO) created for its further development. Utilizing multiple piplines and a wide bandwidth memory connection, the VolumePro500 can process volume data at unprecedently high rates.

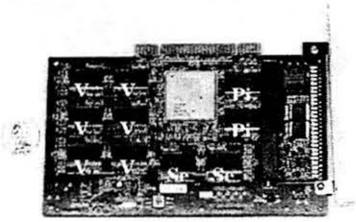
Collaboration: Several applications of this innovative graphics hardware are being investigated by Melco business units. The development of the VolumePro 500 was carried out in close collaboration with MERL, particularly in its early stages.

Future Directions: See the report on the VolumePro 1000.

Contact: Beverly J. Schultz June 22, 2000

Reference: http://www.rtviz.com

VolumePro500-2X



VolumePro500-2X is a new product which just consists of doubling the memory on the existing board. Through the clever use of stacked memory, the team was able to use the same board design and give the customer an alternative that allows them to load bigger images, which is critical as this work continues. Thus, this board allows 256Mbytes of usable memory on the board. It began shipping in December 99,and is the predominant VolumePro offering shipping at the end of the fiscal year.

Background and objectives: The industry wanted to handle larger data sets. VolumePro500 had 128Mbytes of available memory on the board for customers to load volumes. This was not sufficient for many of our early customers. A way to handle larger images was important. This is a key goal of the VolumePro1000, but something we wanted to get to market as fast as possible.

Technical discussion: The senior engineer working on board design continued to scan the literature for new innovations in memory technology. He found references to stacked memory, and if we bought double the memory, it could be taken to this company to be "stacked", giving access to twice as much memory per board.

The results of building a new board with this innovative method of stacking memory chips is so compelling to our customers that within 3 months almost 100% of all board sales were for the 2X board. It has replaced the VolumePro500 in the marketplace.

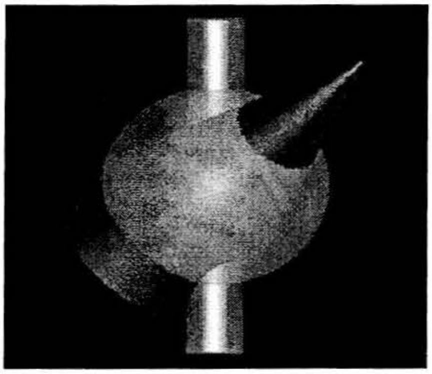
Collaboration: VolumePro500-2X is a project of the Semiconductor Business Group. It is expected that MELCO will use the board in radar applications, medical radiation therapy planning devices, and other applications where visualization is important.

Future Directions: Numerous modifications could be made to the board, as well as the software, to give us additional products that customers would need.

Contact: Beverly J. Schultz June 22, 2000

Reference: http://www.rtviz.com

VolumePro1000



The next product in the VolumePro line will be the VolumePro1000. It will meet key needs of medical and geophysical customers and others who need multiple voxel formats and embedded geometry. The plan is to have the chip run at one billion samples/second to handle large images.

Background and objectives: The product will be a board with software that fits in a PCI slot on an NT machine or in selected Unix systems. The chip capability allows a PCI card with up to 1 GB of memory, 1 billion tri-linearly interpolated, Phong-shaded samples per second (up to 512^3 at 30 fps). It will support embedded opaque and translucent polygon surfaces. The chip will support 8-bit to 32-bit voxels with up to 4 channels and X, Y, and Z supersampling. The innovative pipelines will support either classification before interpolation or interpolation before classification to maximize the capabilities of applications that need such precision. The product has the ability to handle embedded surfaces, multi-channel volume data sets, and both anatomical and functionally segmented data sets.

Technical discussion: The ray casting pipeline has innovative techniques for orthographic projections. The technology architecture allows data transversal, re-sampling using tri-linear interpolation at 1 billlion samples per second, classification by assigning RGBA to each sample, illumination with an estimate of gradients or surface normals from the original data, and finally compositing to blend the samples into pixel color. It can support flexible voxel formats, including RGBalpha voxels, luminance-alpha voxels, and the ability to set categories on a per voxel basis. Highlight is drawing opaque polygons with OpenGL to embed them into a volume.

Key to handling the complex 3D image requirements of the medical and seismic fields is the ability to handle 8-bit, 16-bit, 32-bit voxels and up to four variable-side data channels. To reduce artifacts that are currently present in 3D texture mapping, this chip allows opacity weighting and alpha correction. Four classification look-up tables will be described.. This LSI supports up to 1GB of voxel memory, which is enough to store a 1024x1024x512 volume data set. Performance is achieved with a parallel pipeline architecture and innovative memory controller technology to maximize bandwidth.

.

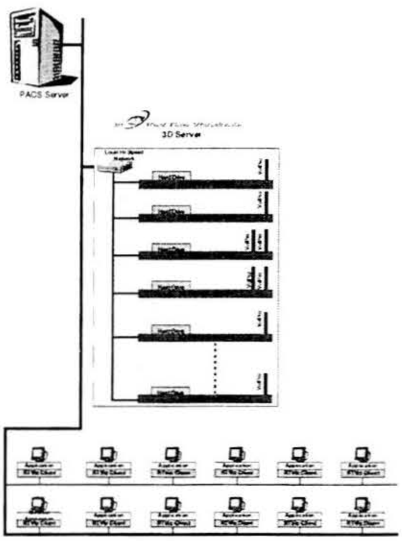
Collaboration: VolumePro1000 is a project of the Semiconductor Business Group. VGO collaborates with MERL on advanced graphics technology.

Future Directions: Future VolumePro will expand data size and flexibility even further.

Contact: Beverly J. Schultz June 22, 2000

Reference: http://www.rtviz.com

3D Web Server



VGO's 3D Web Server consists of a server PC which connects to multiple clients. **3D Web Server Application Project** is the project that the company has instigated in order to optimize the power of our board with a software application that takes advantage of our boards to allow customers to easily do 3D volume visualization. The plan is that a server box will be produced that will have multiple cpu's with multiple VolumePro500 boards in it, software to handle and manipulate the server, and software for efficient handling and rendering of volumes on the server. With the server, clients are available which allow easy access to the server, and can be incorporated into an OEM's application to give them the power of 3D volume rendering in what we expect to previously only be 2D flat images. This project is on target to be displayed at RSNA in November 00 and to be shipping by April 01.

Background and objectives: Hospitals will not put a VolumePro board in every computer in the hospital. However, many doctors would like to walk to the nearest computer and look at 3D visulization of the image that is of interest. The project was created to build a server with numerous clients that could meet such a need.

Technical discussion: The product consists of a set of hardware which will be scalable, with each node being a dual processor motherboard running WinNT with VolumePro. It connects directly to the hospital network or to a PACS server. It is configurable with multiple VP500's per imaging node. The hardware deliverable requires reliability through redundancy and automatic failover. An Admin Manager tracks sessions, software updates and system integrity. A client can connect to the next available image node, and the image nodes can load studies from the DICOM server. Node renders are based on user action. The output image JPEG is compressed and transmitted to the client, which will decompress and display the image.

Collaboration: 3DWeb Server is a project of the Semiconductor Business Group and Corporate R&D. Units within Melco have already shown interest in this product. Collaboration will be required with key partners to fit this product into their applications to enable 3D.

Future Directions: Once the basic product is complete, it is expected that follow-on products will be created that will allow the client to be at a remote site, rather than in the hospital.

Contact: Beverly J. Schultz June 22, 2000

Reference: http://www.rtviz.com

MERL Cambridge Research

MERL Cambridge Research pursues application-motivated basic research in computer science. We develop new technologies looking five years and more into the future; however, our efforts are directed toward applications of practical significance. By connecting to our sister organizations in MERL and MELCO, we forge collaborations that carry our technologies into products.

Good communication is, we believe, the key to successful research. The best ideas are born, and mature most quickly, when critically examined and refined by many minds. Toward this end, MERL researchers are encouraged to work together with each other and with researchers at other institutions. Further, we are an active member of the research community, publishing the work we do as quickly as possible.

To allow rapid response to opportunities, MERL Cambridge Research is organized as a flexible community of researchers without internal divisions. Our goal is to support a continually changing mix of individual explorations and group projects, where promising individual efforts can easily grow into projects and projects can easily disband at the end of their natural lives. Each researcher is typically involved in both individual explorations and related group projects.

As documented in the following pages, MERL Cambridge Research currently focuses on three intertwined themes: computer vision, computer graphics, and human/computer interaction.

In computer vision, we focus on real-time processing; exploring ways that computer vision can be used as an input modality. An interesting sub-theme is computer learning, which is often required to acquire the base data needed for visual recognition and processing.

In computer graphics, we focus on real time processing and volume graphics. Our work on volume rendering led to the creation of MERL Concord and the real time volume-rendering chip they developed.

In human/computer interaction, our work includes improved computer interfaces with intelligent agents, graphical interfaces, and human/computer collaboration. We are also interested in ubiquitous computing---the rapidly growing trend toward embedding computers in special purpose devices rather than using them in a general-purpose form

Located in Cambridge, Massachusetts, adjacent to the Massachusetts Institute of Technology, the laboratory currently consists of 25 researchers with 5 administrative and support staff members. The expertise of the researchers ranges from mathematics to computer software to the social sciences, with some of the best work arising where these different disciplines interact. The permanent staff is enriched by an active program of student internships hosting approximately 40 students per year for an average of three to four months each.

Successful basic research requires a supportive environment, open communication, access to real-world problems, and a long-term perspective. We are committed to providing all of these elements in full measure at MERL Cambridge Research.

Technical Staff

David B. Anderson

M.Sc., Carnegie Mellon University, 1985

David Anderson is most interested in ideas that combine technical innovation, artistic vision, and business opportunity. Before joining MERL Cambridge Research, he worked on the Mach and Andrew systems projects at CMU, and taught at the Pennsylvania Governor's School for the Sciences. At MERL Cambridge Research, he has been part of creating the Diamond Park shared virtual world, the SPLINE platform for distributed multimedia, and Open Community. More recently, he has been developing applications that use tangible and perceptual interfaces. David Anderson also serves on the board of directors of the Web3D Consortium.

Paul Beardsley

Ph.D., Oxford University, 1992

Paul Beardsley is a computer vision researcher. He obtained a DPhil from Oxford for work on motion recovery and camera calibration using projective geometry. A post-doc followed this on navigation through unknown environments, sensing the environment with uncalibrated cameras undergoing unknown motion. His recent work has been on face tracking, and he is currently doing advanced development work on a system for extracting 3D face models using stereo cameras.

Matthew Brand

Ph.D., Northwestern University, 1994

Matthew Brand's work focuses on learning how the world behaves from sensory data. One goal is to make machines that steadily become better at interpreting, assisting, and even mimicking human activity. Recent results include an entropy optimization framework for learning and methods for synthesizing high-quality virtual data from probability models. These techniques are being demonstrated in "digital puppets" -- systems that learn to synthesize realistic, expressive human behavior. In 1999 Brand was named one of 100 top innovators of his generation by Technology Review.

Michael A. Casey

Ph.D., Massachusetts Institute of Technology, 1998

Michael Casey's research focus is in developing novel signal processing and pattern recognition techniques for analysis and synthesis of audio. For his thesis Michael developed a new audio analysis technique based on independent components analysis that makes possible perceptually salient feature decomposition of textured and noise-based sounds and sound scenes. As an intern at MERL Cambridge Research before completing his degree, he also participated in the development of audio software for SPLINE and produced the sound design for Diamond Park, a large-scale multi-participant virtual environment.

William T. Freeman

Ph.D., Massachusetts Institute of Technology, 1992

Freeman, William T. is a senior research scientist at MERL Cambridge Research, where he studies machine learning applied to computer vision and interactive applications of computer vision. As part of his doctoral work, he developed "steerable filters", a class of oriented filters useful in image processing and computer vision. In 1997 he received the Outstanding Paper prize at the Conference on Computer Vision and Pattern Recognition for work on applying bilinear models to separate "style and content". During 1987-88 he was a Foreign Expert at the Taiyuan University of Technology, Taiyuan, Shanxi, China

Sarah Frisken

Ph.D., Carnegie Mellon University, 1991

Sarah Frisken has research interests in computer graphics, volume visualization and physically based modeling. She has led a team of researchers and students to build a knee arthroscopy simulator that incorporates high-quality rendering, haptic feedback and physical modeling to simulate interactions between surgical tools and a computer model derived from 3D MRI data. Her current work is with Adaptively Sampled Distance Fields, a general representation of shape for computer graphics, which provides intuitive manipulation, and editing of smooth surfaces with fine detail. Applications include electronic sculpting, volumetric effects for rendering, color gamut representation, path planning for CNC milling, and rapid prototyping.

Ray Jones

B.Sc., University of Utah, 1994

Ray Jones's research interests include real-time rendering methods, efficient data representations for graphics applications, and graphics modeling tools. He joined MERL Cambridge Reserach in 1999. His previous work included developing a 3D paint system and optimization work on the SunPHIGS structure walker. His current research is on high-quality real-time rendering systems.

Darren Leigh

Ph.D., Harvard University, 1998

Darren Leigh's research interests range from electronic hardware and communications to operating systems and signal processing. Before coming to MERL Cambridge Research he worked on the Harvard University/Planetary Society Billion-channel ExtraTerrestrial Assay (Project BETA), a search for microwave signals from extraterrestrial civilizations. Other previous research includes 3D microscopic scanning, desktop manufacturing and network architectures for multimedia. His current research includes the Personal Eyewitness and interfacing and applications of the M32R/D and Artificial Retina chips.

Neal Lesh

Ph.D., University of Washington, 1998

Neal Lesh has been studying techniques for improving human-computer, collaborative problem solving. Before coming to MERL Cambridge Research, he completed a thesis on scalable and adaptive goal recognition at the University of Washington, and worked as a postdoc with James Allen at the University of Rochester on the TRIPS collaborative planning project. At MERL Cambridge Research, he is working on systems for interactively solving hard combinatorial optimization problems, as well as on technology for building collaborative interface agents, called COLLAGEN. Recently he has begun working on techniques to allow people to share experiences using digital media.

Joe Marks

Ph.D., Harvard University, 1991

Joe Marks worked previously at Bolt Beranek and Newman and at Digital's Cambridge Research Laboratory. His main areas of interest are computer graphics, user interfaces, and heuristic optimization. His current projects at MERL Cambridge Research concern the development of novel tangible and perceptual user interfaces for design tasks. Dr. Marks also has a strong interest in teaching. He was an adjunct lecturer in the Division of Engineering and Applied Sciences at Harvard University from 1991 to 1996, and currently teaches two courses at Harvard Extension School.

Brian Mirtich

Ph.D., University of California at Berkeley, 1996

Brian Mirtich's research interests lie in the broad area of scientific computing. He is interested in the algorithmic and mathematical techniques needed to solve complex problems. Rigid body dynamic simulation has been a major focus of his research; recent work explores how to efficiently simulate very large systems. Computational geometry, particularly collision detection algorithms, is another area of emphasis. Current work includes topics from combinatorial optimization and representation of three-dimensional data. He has published widely used code for various applications in simulation and modeling.

Baback Moghaddam

Ph.D., Massachusetts Institute of Technology, 1997

Baback Moghaddam's research interests are in computational vision and image processing with special focus on probabilistic visual learning, statistical and neural network modeling and pattern recognition. Prior to coming to MERL Cambridge Reserach, Dr. Moghaddam was at the Vision and Modeling Group at the MIT Media Laboratory. As part of his dissertation, he developed an automatic face recognition system that won ARPA's 1996 "FERET" face recognition competition. Past research includes fractal image compression, segmentation and analysis of SAR and IR imagery as well as designing a zero-gravity experiment for laser annealing of amorphous silicon flown aboard the U.S. space shuttle in 1990.

Ron Perry

B.Sc., Bucknell University, 1981

Ron Perry joined MERL Cambridge Research as a Research Scientist in 1998. Prior to that, he was a consulting engineer at DEC developing a three-dimensional rendering ASIC called Neon. Ron has consulted for many companies including Kodak, Adobe, Quark, and Apple over the last 18 years, developing software and hardware products in the areas of computer graphics, imaging, color, and desktop publishing. Some key product developments include the color engines for QuarkXPress, Adobe PhotoShop, Adobe Illustrator, and Windows 95/98 as well as the Atex Display and Pagination system which is used by most major metropolitan newspapers in the world to paginate, display, and print their publications.

Hanspeter Pfister

Ph.D., State University of New York at Stony Brook, 1996

Hanspeter Pfister joined MERL Cambridge Research as a Research Scientist in 1996. He is the chief architect of VolumePro, Mitsubishi Electric's real-time volume rendering system for PC-class computers. His research interests include computer graphics, scientific visualization, computer architecture, and VLSI design. Hanspeter Pfister received his Ph.D. in Computer Science in 1996 from the State University of New York at Stony Brook. In his doctoral research he developed Cube-4, a scalable architecture for real-time volume rendering. He received his Dipl.-Ing. degree in electrical engineering from the Department of Electrical Engineering at the Swiss Federal Institute of Technology (ETH) Zurich in 1991. He is a member of the ACM, IEEE, the IEEE Computer Society, and the Eurographics Association.

Stanley W. Pozerski, Jr.

BA Computer Systems, Daniel Webster College 1987

Stan's interests have followed the application of computers to a variety of manufacturing tasks including using PDP-11's to demonstrate control of multiple reactor chemical processes, using personal computers for production testing, manufacturing chemicals and controlling multi-axis rotary assembly machines. More recently, Stan has been Systems Administrator of a CIM system for a semiconductor facility performing shop floor scheduling, data collection, and process monitoring. Currently, Stan supports Windows and Linux clients and servers, networking, and the wide variety of PC applications used at MERL Cambridge Research.

Charles Rich

Ph.D., Massachusetts Institute of Technology, 1980

The thread connecting all of Dr. Rich's research has been to make interacting with a computer more like interacting with a person. As a founder and director of the Programmer's Apprentice project at the MIT Artificial Intelligence Laboratory in the 1980s, he pioneered research on intelligent assistants for software engineers. Dr. Rich joined MERL Cambridge Research in 1991 as a founding member of the research staff. For the past several years, he has been working on a technology for building collaborative interface agents, called COLLAGEN, which is based on a theory of human collaborative discourse. Dr. Rich is a Fellow and past Councilor of the American Association for Artificial Intelligence.

Alyn Rockwood
Ph.D., Cambridge University, 1986

Former professor of computer science at Arizona State University (90-99) and industrial researcher for 13 years at Evans and Sutherland and SGI, interests and publications include scientific computing, scientific visualization, mechanical CAD, geometric design, computer graphics and image processing. Currently, he is investigating fundamental structures and mathematical underpinnings for computer graphics. He is an author of 3 technical books and was SIGGRAPH99 Papers' Chair.

Chia Shen
Ph.D., University of Massachusetts, 1992

Chia Shen's current research is in distributed real-time and multimedia systems. She is particularly interested in the non-traditional use of standard high-speed networks, such as ATM, for distributed industrial control applications and distributed multimedia environments. Her work involves the design of middleware algorithms and protocols. In her Ph.D. research on the Spring kernel, she worked on issues in multiprocessor real-time operating systems including on-line scheduling, dispatching, resource reclamation, and predictable synchronization mechanisms.

Candace L. Sidner
Ph.D., Massachusetts Institute of Technology, 1979

Candy Sidner is an expert in user interfaces, especially those involving speech and natural language understanding, and human and machine collaboration. Before coming to MERL Cambridge Research, she had been a member of the research staff at Bolt Beranek Newman, Inc., Digital Equipment Corp., and Lotus Development Corp., and a visiting scientist at Harvard University. She is currently working, among other things, on how to apply speech understanding technology to collaborative interface agents in the COLLAGEN project. Dr. Sidner is currently Chair of the 2001 International Conference on Intelligent User Interfaces and is a past President of the Association for Computational Linguistics. She is also a Fellow and past Councilor of the American Association for Artificial Intelligence.

Carol Strohecker
Ph.D., Massachusetts Institute of Technology, 1991

Carol Strohecker is concerned with how people learn and how objects, artifacts, and technologies can facilitate learning. Research interests include interactive narrative and environments that support constructive activities in complementary physical and virtual domains. Dr. Strohecker is a Presidential Nominee on the MIT Corporation Visiting Committee for Media Arts and Sciences. She is an advisor for the Making Models program at Boston's Museum of Science, and has been a Fellow of the Harvard University Graduate School of Design, the Massachusetts Council for the Arts and Humanities, and the U.S. National Endowment for the Arts.

Jeroen van Baar

M.Sc., Delft University of Technology, 1998

Jeroen van Baar's interest is on computer graphics, scientific visualization and currently more specific on interactive rendering techniques and data representations for such techniques. He has been an intern at MERL Cambridge Research in 1997, 1998 and 1999 and has joined MERL Cambridge Research as a full-time member in 2000. Jeroen is currently working on a point-based rendering technique called surfels.

Jonathan S. Yedidia

Ph.D., Princeton University, 1990

Jonathan Yedidia's graduate work at Princeton and post-doctoral work at Harvard's Society of Fellows focused on theoretical condensed-matter physics, particularly the statistical mechanics of systems with "quenched" disorder. He was a professional chess player and teacher from 1993 to 1997. He then joined the internet startup company Viaweb, where he worked on a shopping search engine that has since become shopping.yahoo.com. Dr. Yedidia is interested in the application of statistical methods to inference and learning.

William S. Yerazunis

Ph.D., Rensselaer Polytechnic Institute, 1987

William Yerazunis has worked in a number of fields including: optics and signal processing (for General Electric's jet engine manufacturing); computer graphics (at Rensselaer's Center for Interactive Computer Graphics); artificial intelligence and parallel symbolic computation (for DEC's OPS5, XCON, and the successor products such as RuleWorks); radioastronomy and exobiology (at Harvard University), and transplant immunology (for the American Red Cross). He holds 15 U.S. patents.

Recent Major Publications

Anderson, David., Frankel, Jamie., Marks, Joe., Agarwala, Aseem, Beardsley, Ryan., Hodgins, Jessica., Leigh, Darren., Ryall, Kathy., Sullivan, Ed., Yedidia, Jonathan., "Tangible Interaction + Graphical Interpretation: A New Approach to Intuitive 3D Modeling," *to appear in Proceedings of Siggraph 2000*, ACM, July 2000 (TR2000-13).

Brand, Matthew., Hertzmann, Aaron "Style Machines." *to appear Proceedings SIGGRAPH 2000, ACM, July 2000.* TR2000-14

Frisken, Sarah F., Perry, Ronald N., Rockwood, Alyn P., Jones, Thouis R. ,"Adaptively Sampled Distance Fields: A General Representation of Shape for Computer Graphics", *to appear Proceedings Siggraph '00, ACM, July 2000, (TR2000-15)*

Mirtich, Brian, "Timewarp Rigid Body Simulation", *to appear SIGGRAPH 00 Conference Proceedings,* July 2000, Addison Wesley.

Pfister, Hanspeter, Zwicker, Matthias., van Baar, Joeron, Gross, Marcus "Surfels – Surface Elements as Rendering Primitives", *to appear in Proceedings of SIGGRAPH 2000*, TR2000-10.

Anderson, David., Anderson, Emily., Lesh, Neal., Marks, Joe., Mirtich, Brian, Ratajczak, David., Ryall, Kathy., "Human-Guided Simple Search" *to appear in the Seventeenth National Conference on Artificial Intelligence (AAAI-2000),* July 30-August 3, 2000, TR2000-16.

Bhalerao, A. Pfister, Hanspeter, Halle, M., and Kikinis, R., "Fast Re-Rendering of Volume and Surface Graphics by Depth, Color, and Opacity Buffering", *to appear Journal of Medical Image Analysis , in May 2000.*

Matthew Brand and Vera Kettnaker. "Inferring hidden state in video." *to appear IEEE Transactions on Pattern Analysis and Machine Intelligence*, volume 22. TR99-22.

Freeman, William T., Pasztor, Egon C., Carmichael, Owen T.,"Learning Low-Level Vision", *accepted for publication in International Journal of Computer Vision*, 2000, TR2000-05.

Freeman, William T., Weiss, Yair, "On the fixed points of the max-product algorithm", *accepted IEEE Transactions on Information Theory*, 2000, TR99-39

Jones, Thouis R., Perry, Ronald N., "Antialiasing with Line Samples", *to appear Rendering Techniques 2000 Proceeding Seventh Eurographics Workshop on Rendering.*

Weiss, Yair and Freeman, William T., "Correctness of belief propagation in Gaussian graphical models of arbitrary topology" *to appear in Advances in Neural Information Processing Systems 12*}, edited by S. A. Solla, T. K. Leen, and K-R. Muller, 2000, TR99-38.

Anderson, David, Anderson, Emily, Lesh, Neal, Marks, Joe, Mirtich, Brian, Ratajczak, David, Ryall, Kathy, "Human-Guided Simple Search" *Seventeenth National Conference on Artificial Intelligence*, July 30-August 3, 2000, TR2000-16

Anderson, Emily, Frankel, Jamie, Marks, Joe, Agarwala, Aseem, Beardsley, Paul, Hodgins, Jessica, Leigh, Darren, Ryall, Kathy, Sullivan, Edward, Yedidia, Jonathan, "Tangible Interaction + Graphical Interpretation: A New Approach to Intuitive 3-D Modeling," *Proceedings of SIGGRAPH 2000*, ACM, July 2000 TR2000-13

Frisken-Gibson, Sarah , Perry, Ronald N., Rockwood, Alyn P., Jones, Thouis R., "Adaptively Sampled Distance Fields: A General Representation of Shape for Computer Graphics", *Proceedings SIGGRAPH 2000*, ACM, July 2000, TR2000-15

Mirtich, Brian, "Timewarp Rigid Body Simulation", *SIGGRAPH 2000 Conference Proceedings*, July 2000

DeMaso, D. R., Gonzalez-Heydrich, J., Erickson, J. D., Grimes, V. P., Strohecker, Carol "The Experience Journal: A Computer-Based Intervention for Families Facing Congenital Heart Disease", *J AM Acad Child Adolesc Psychiatry*, June 2000, TR2000-12

Bhalerao, A., Pfister, Hanspeter, Halle, M. and Kikinis, R., "Fast Re-Rendering of Volume and Surface Graphics by Depth, Color, and Opacity Buffering", *Journal of Medical Image Analysis*, May 2000

Strohecker, Carol, and Slaughter, Adrienne H. "Kits for Learning and a Kit for Kitmaking", *Proceedings of CHI'2000,* ACM Press, April 2000, TR2000-02

Freeman, William T. and Pasztor, Egon, "Markov Networks for Super-Resolution", *Proceedings 34th Annual Conference on Information Sciences and Systems*, March 2000, TR2000-08

Freeman, William T., "Biometrics: personal identification in a networked society", *Pattern Analysis and Applications*, March 2000, TR2000-07

Moghaddam, Baback, and Yang, M-H., "Gender Classification with Support Vector Machines," in *Proceedings of the 4th IEEE International Conference on Face and Gesture Recognition,* Grenoble, France, March 2000, TR2000-01

Strohecker, Carol, and Barros, B. "Make Way for WayMaker", *Presence: Teleoperators and Virtual Environments*, v9 n1, pp.97-107, February 2000, TR97-07a

Weiss, Y., and Freeman, William T., "Correctness of Belief Propagation in Gaussian Graphical Models of Arbitrary Topology" *Advances in Neural Information Processing Systems 12*, 2000. TR99-38

Freeman, William T., Pasztor, Egon, Carmichael, Owen, "Learning Low-Level Vision", *International Journal of Computer Vision*, 2000, TR2000-05

Freeman, William T., Weiss, Y., "On the Fixed Points of the Max-Product Algorithm", *IEEE Transactions on Information Theory*, 2000, TR99-39

Howe, N., Leventon, Michael, Freeman, William T., "Bayesian Reconstruction of 3-D Human Motion from Single-Camera Video", *Advances in Neural Information Processing Systems 12*, 2000, TR99-37

Tenenbaum, J. and Freeman, William T., "Separating Style and Content with Bilinear Models", *Neural Computation,* v12 n6, 2000, TR99-04

Jones, Thouis R., and Perry, Ronald N., "Antialiasing with Line Samples", Rendering Techniques 2000 *Proceeding Seventh Eurographics Workshop on Rendering.*

Pfister, Hanspeter, Zwicker, Matthias, van Baar, Jeron, Gross, Marcus, "Surfels- Surface Elements as Rendering Primitives", *Proceedings of SIGGRAPH 2000*, TR2000-10

Strohecker, Carol "Cognitive Zoom: From Object to Path and Back Again". *Spatial Cognition II, Springer-Verlag,* 2000, TR2000-04

Yedidia, Jonathan S., Freeman, William T., Weiss, Yair, "Generalized Belief Propagation", *Neural Information Processing Systems*, 2000, TR2000-26

Weiss, Y., and Freeman, William T., "Correctness of Belief Propagation in Gaussian Graphical Models of Arbitrary Topology", *Neural Computation*, TR99-38

Kage, H., Freeman, William T., Miyake, Y., Funatsu, E., Tanaka, K., Kyuma, K., "Artificial Retina Chips as On-Chip Image Processors and Gesture-Oriented Interfaces", *Optical Engineering*, v38 n12, pp. 1979-1988. December, 1999, TR2000-06

Freeman, William T., Beardsley, Paul A., Kage, H., Tanaka, K., Kyuman, C., Weissman, C., "Computer Vision for Computer Interaction", *SIGGRAPH Computer Graphics Magazine*, November, 1999, TR99-36

Moghaddam, Baback, Biermann H., Margaritis D., "Regions-of-Interest and Spatial Layout in Content-based Image Retrieval," *European Workshop on Content-Based Multimedia Indexing*, Toulouse, France, October 1999, TR99-10

Golding, Andrew, and Lesh, Neal, "Indoor Navigation Using a Diverse Set of Cheap, Wearable Sensors", *The Third International Symposium on Wearable Computers*, October 18-19, 1999, San Francisco, pp. 29-36, TR99-32

Ray, H., Pfister, Hanspeter, Silver, D., Cook, T.A., "Ray-Casting Architectures for Volume Visualization", *IEEE Transactions on Visualization and Computer Graphics*, pp. 210-223, v5 n3, July/September 1999, TR99-17

Moghaddam, Baback, "Principal Manifolds and Bayesian Subspaces for Visual Recognition," *Proceedings of the 7th International Conference on Computer Vision,* Corfu, Greece, September 1999, TR99-35

Pfister, Hanspeter, Hardenberg, J., Knittel, J., Lauer, H., Seiler, L., "The VolumePro Real-Time Ray-Casting System", *Proceedings of SIGGRAPH 1999*, pp. 251-260, Los Angeles, CA, August 1999

Lesh, Neal, Zaki, M., Ogihara, M., "Mining features for Sequence Classification", *5th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Diego, CA, August 1999, TR98-22

Lesh, Neal, and Allen, J., "Simulation-based Inference for Plan Monitoring", *Sixteenth National Conference on Artificial Intelligence*, Orlando, July 18-22, 1999 pp.358-365, TR99-06

Shen, Chia, Gonzalez, O., Ramamritham, K., Mizunuma, I., "User Level Scheduling of Communicating Real-Time Tasks", *IEEE 1999 Real-Time Technology and Applications Symposium*, June 2-4, 1999, Vancouver, British Columbia, Canada, pp. 164-175, TR99-21

Moghaddam, Baback and Pentland A., "Bayesian Image Retrieval in Biometric Databases," *IEEE International Conference on Multimedia Computing and Systems,* Florence, Italy, June, 1999, TR99-13

Moghaddam, Baback, Henning B., and Dimitris M., "Defining Image Content with Multiple Regions-of-Interest", *Proceedings of IEEE Workshop on Content-based Access of Image and Video Libraries*, June 1999, pp. 89-93, TR99-10

Lesh, Neal, Rich, Charles, Sidner, Candace L., "Using Plan Recognition in Human-Computer Collaboration", in *Proceedings of Seventh International Conference on User Modeling*, Banff, Canada, June 1999, pp. 23-32, TR98-23

Strohecker, Carol, "Toward a Developmental Image of the City: Design Through Visual, Spatial, and Mathematical Reasoning", in *Proceedings of Visual and Spatial Reasoning in Design*, June, 1999, pp.33-50, TR99-07

Yerazunis, William S., Leigh, Darren L., Freeman, William T., and Bardsley, R. S., "An Inexpensive, All Solid-state Video and Data Recorder for Accident Reconstruction", *1999 SAE International Congress and Exposition*, Detroit. SAE Technical Paper #1999-10-1299, March 1999, TR99-29

Brand, Matthew, "Pattern Discovery via Entropy Minimization", *Proceedings of Uncertainty '99*, Society of Artificial Intelligence and Statistics, January 1999, TR98-21

Moghaddam, Baback, Jebara,T., and Pentland, A., "Bayesian Modeling of Facial Similarity", *Advances in Neural Information Processing Systems 11*, January 1999, pp. 910-916, TR99-13

Pfister, Hanspeter, "Architectures for Real-Time Volume Rendering", *Future Generation Computer Systems*, v15, 1999, pp. 1-9, TR99-18

Freeman, William T. and Pasztor, Egon, "Learning to Estimate Scenes from Images", *Advances in Neural Information Processing Systems*, v11, 1999, TR2000-08

Freeman, William T. and Pasztor, Egon, "Learning Low-Level Vision", *IEEE International Conference on Computer Vision*, Corfu, Greece, 1999, TR99-12

Frisken-Gibson, Sarah, "Using Linked Volumes to Model Object Collisions, Deformation, Cutting, Carving, and Joining", *IEEE Transactions on Visualization and Computer Graphics*, v5 n4, 1999

Hornby, Greg., and Mirtich, Brian "Diffuse versus True Coevolution in a Physics-based World", *Proceedings of 1999 Genetic and Evolutionary Computation Conference*, pp. 1305-1312, TR98-11

Strohecker, Carol "Construction Kits as Learning Environments", Proceedings of the International Conference on Multimedia Computing and Systems, IEEE Computer Society, 1999, II, pp.1030-1031, TR99-31

Strohecker, Carol, Barros, B., and Slaughter, Adrienne H., "Mapping Psychological and Virtual Spaces", *Online Proceedings of Design Computing on the Net*, November/December 1998. TR98-12

Brand, Matthew, and Shan, K., "Voice-Driven Animation", *Workshop on Perceptual User Interfaces*, November 1998, TR98-20

Beardsley, Paul A., "A Qualitive Approach to Classifying Head and Eye Pose", *IEEE Workshop on Applications of Computer Vision*, October 1998, pp.208-213, TR98-10

Frisken-Gibson, Sarah F., "Using Distance Maps for Accurate Surface Representation in Sampled Volumes", *Proceedings 1998 IEEE Symposium on Volume Visualization*, October 1998, pp. 23-30, TR99-25

Frisken-Gibson, Sarah F., "Constrained Elastic SurfaceNets: Generating Smooth Models from Binary Segmented Data", *Medical Image Computing and Computer-Assisted Interventions-MICCAI '98*, Lecture Notes in Computer Science 1496, October 1998, pp. 888-898, TR99-24

Schill, Marcus, Frisken-Gibson, Sarah F., Bender, H., Manner, R., "Biomechanical Simulation of the Vitreous Humor in the Eye Using an Enhanced ChainMail Algorithm", *Medical Image Computing and Computer-Assisted Interventions*, Lecture Notes in Computer Science 1496, October 1998, pp. 679-687, TR99-27

Brand, Matthew, "Structure Discovery in Conditional Probability Distributions via an Entropic Estimator", *Neural Computation*, August 1998, TR98-18

Mirtich, Brian, "*V-Clip:* Fast and Robust Polyhedral Collision Detection", *ACM Transactions on Graphics*, v17 n3, July 1998, pp. 177-208, TR97-05

Freeman, William T., Anderson, David, Beardsley, Paul, Dodge, C., Kage, H., Kyuman, K., Miyake, Y., Rother, M., Tanaka, K., Weissman, C., Yerazunis, William T., "Computer Vision for Interactive Computer Graphics", *IEEE Computer Graphics and Applications*, May-June, 1998, pp. 42-53, TR99-02

Frisken-Gibson, Sarah, Fyock, C., Grimson, E., Kanade, T., Kikinis, R., Lauer, Hugh, McKenzie, Neil, Mor, A., Nakajima, S., Ohkami, H., Osborne, R., Samosky, J., Sawada, A., "Simulating Surgery Using Volumetric Object Representations, Real-Time Volume Rendering and Haptic Feedback", *Medical Image Analysis*, February 1998, pp. 121-132, TR97-02

Freeman, William T. and Viola, P., "Bayesian Model of Surface Perception", *Neural Information Processing Systems*, v10, 1998, TR98-05

Perona, P. and Freeman, William T., "A Factorization Approach to Grouping", *European Conference on Computer Vision*, 1998, TR99-03

Mirtich, Brian, "Efficient Algorithms for Two-Phase Collision Detection", *Practical Motion Planning in Robotics: Current Approaches and Future Directions*, pp. 203-223, 1998 TR97-03

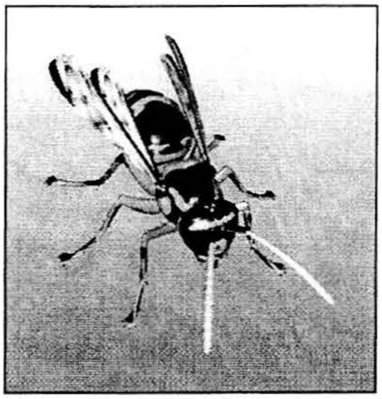
Moghaddam, Baback and Pentland, A., "Beyond Euclidean Eigenspaces: Bayesian Matching for Visual Recognition", *Face Recognition: From Theories to Applications*, NATO ASI Series F vol. 163, 1998 TR98-08

Rich, Charles and Sidner, Candace L. "COLLAGEN: A Collaboration Manager for Software Interface Agents", *User Modeling and User-Adapted Interaction, Special Issue on Computational Models for Mixed Initiative Interaction*, v8 n3/4, 1998, pp. 315-350, TR97-21

Bers, M. U., Ackermann, E., Cassell, J., Donegan, J., Gonzalez-Heydrich, J., DeMaso, D. R., Strohecker, Carol, Lualdi, S., Bromley, D., and Karlin, J. "Interactive storytelling environments: Coping with cardiac illness at Boston's Children's Hospital". *Proceedings of CHI '98, Human Factors in Computing Systems*, pp. 603-610, Association for Computing Machinery, TR98-09

Project Descriptions

SURFELS - Surface Elements as Rendering Primitives



See Color Plate #1

This project is motivated by the desire to render complex three dimensional objects as efficiently as possible. The fundamental approach of this research is to use *surfels* (or surface elements) as an alternative display primitive. Objects are represented as a dense set of surface point samples. Surfel attributes comprise depth, texture color, normal, and others. As a pre-process, an octree-based surfel representation of a geometric object is computed. During rendering, a hierarchical forward warping algorithm projects surfels to the screen. Surfel objects offer complex shape, low rendering cost and high image quality, which makes them specifically suited for low-cost, real-time graphics, such as games.

surfels: surface elements

Background and objectives: The core objective of this project is to *combine volume, image, and geometry rendering* in one unified rendering pipeline. We have developed several new algorithms for surfel generation and rendering and we implemented a prototype surfel rendering system. We have applied for 4 patents in 1999 and are currently applying for 4 additional patents. We are currently exploring hardware implementations of the surfel pipeline.

Technical discussion: We propose a new method of rendering objects with rich shapes and textures at interactive frame rates. Our rendering architecture is based on simple surface elements (*surfels*) as rendering primitives. Surfels are point samples of a graphics model. In a preprocessing step, we sample the surfaces of complex geometric models along three orthographic views. At the same time, we perform computation-intensive calculations such as texture, bump, or displacement mapping. By moving rasterization and texturing from the core rendering pipeline to the preprocessing step, we dramatically reduce the rendering cost. From a modeling point of view, the surfel representation provides a mere discretization of the geometry and hence reduces the object representation to the essentials needed for rendering. In a sense, a surfel relates to what we call the *''lingua franca''* of rendering. Surfels will allow us to combine image-based rendering, adaptive curved-surface tessellation, constructive solid geometry, unstructured volume data, particle systems, and so on in one common rendering framework.

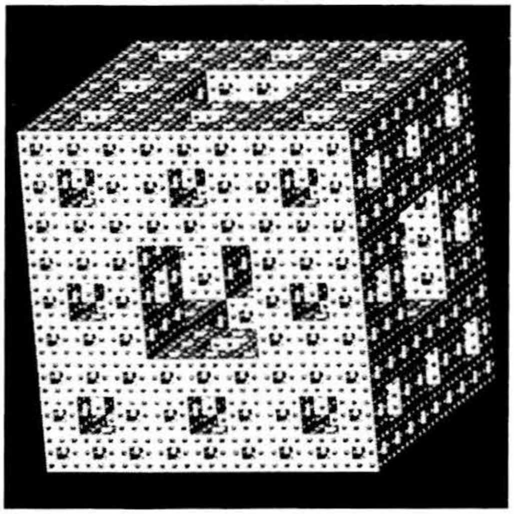
Collaboration: This project currently involves researchers from MERL Cambridge Research and the Swiss Federal Institute of Technology (ETH) in Zürich, Switzerland.

Future Directions: We are improving the rendering pipeline, adding support for volumes and image-based data. We are also developing a surfel rendering architecture amenable for hardware implementation.

Contact: Hanspeter Pfister

February 8, 2000

Adaptively Sampled Distance Fields: A General Representation of Shape for Computer Graphics



Adaptively Sampled Distance Fields (ADFs) are a new representation of shape that allows efficient and accurate representation of both smoothly curved surfaces and surfaces with fine detail.

ADFs are a simple yet consolidating form which can be used to represent an extensive variety of graphical shapes. ADFs support a broad range of processing operations such as rendering, sculpting, level-of-detail management, surface offsetting, collision detection, and color gamut correction

Two applications of ADFs are sculpting fine detail onto objects, and path planning and tool selection for automated machining.

ADF: Adaptively Sampled Distance Field

Background and objectives: A distance field is a scalar field that specifies the distance to a shape, where the distance may be signed to distinguish between the inside and outside of the shape. The distance field is an effective representation of shape and an alternative to surface representations such as polygons or NURBS. However, regularly sampled distance fields have drawbacks because of their size and limited resolution. Because fine detail requires dense sampling, immense volumes are needed to accurately represent classical distance fields with regular sampling when any fine detail is present, even when the fine detail occupies only a small fraction of the volume.

Technical discussion: To overcome the limitations of regularly sampled distance fields, ADFs use adaptive, detail-directed sampling, with high sampling rates in regions where the distance field contains fine detail and low sampling rates where the field varies smoothly. This sampling permits arbitrary accuracy in the reconstructed field together with efficient memory usage. To process the adaptively sampled data efficiently, ADFs store the sampled data in a hierarchy. The combination of detail-directed sampling and the use of a spatial hierarchy for data storage allows ADFs to represent complex shapes to precisely while permitting efficient processing.

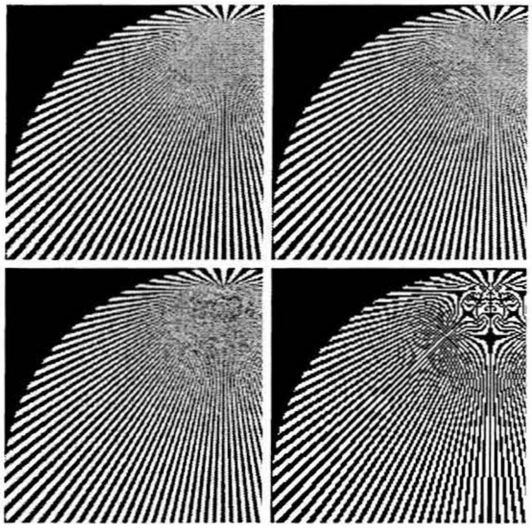
Collaboration: Currently, this work is MERL Cambridge Research-internal but several collaborations with universities and other research groups are being explored.

Future Directions: We continue to investigate fundamental algorithms and methods for ADFs. We are completing an ADF library for limited distribution to test the application of ADFs to several areas including digital sculpting, CAD/CAM, and 3D scanning.

Contacts: Sarah Frisken, Ron Perry, Alyn Rockwood, Ray Jones June 1, 2000

Reference: http://www.merl.com/projects/adfs

Line Sampling



Clockwise from upper left: supersampling (256 samples per pixel), line sampling, regular sampling (1 sample per pixel), supersampling (16 samples per pixel).

Images produced by rendering three-dimensional graphics suffer from a problem known as aliasing. Aliases appear as distracting artifacts, such as jagged edges on polygons and lines. High quality rendering systems use antialiasing to prevent these artifacts.

Line sampling is a new method of antialiasing that gives results comparable to analytic antialiasing without the computational expense. Line sampling does not take much more computation than supersampling, but is specifically designed to avoid the failures that make supersampling unsuitable for high quality applications.

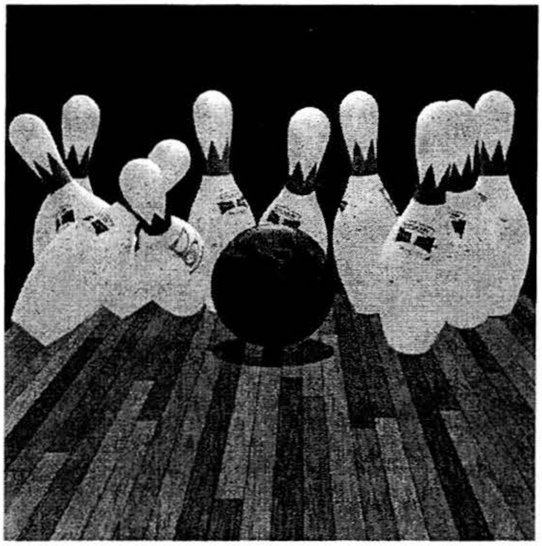
Background and objectives: The recent trend in consumer graphics has been towards higher quality images. Recently, these systems have also started to provide full scene antialiasing. This antialiasing is implemented as a simple form of supersampling, which handles simple cases well but performs poorly on some common scene elements. Consequently, new antialiasing methods will have to be utilitized. In this project, we have developed a new antialiasing method, line sampling, that combines the best qualities of analytic antialiasing with the efficiency of supersampling. Line sampling is specifically designed to perform well on the failure cases of supersampling without incurring the cost of fully analytic methods. The central advancement is the use of one-dimensional samples, rather than the more common zero-dimensional (point) samples used in supersampling or the expensive two-dimensional (area) samples used in analytic methods.

Technical discussion: Antialiasing is, on a theoretical level, the calculation of an integral, the convolution of the graphical scene with some bandlimiting filter. Analytic methods calculate the integral exactly, but the need to calculate the two-dimensional solution to the hidden surface problem drives up their cost and complexity. Supersampling methods use several disjoint point samples to approximate the integral. The hidden surface problem is greatly simplified for point samples, so the cost is much lower. Unfortunately, the point samples provide a discrete approximation of the integral. This causes the value of pixels to jump when objects move across the screen. These jumps appear as twinkling edges in the animation. Line sampling avoids these temporal artifacts by sampling along a one-dimensional line segment. Since this is a continuous sampling, as objects move their edges change color smoothly, removing the twinkling effect. The hidden surface problem is not very complicated, and the rendering cost of line sampling is not much more than that of supersampling. Line sampling produces images almost indistinguishable from those produced by analytic methods.

Future Directions: We plan to build a real-time renderer utilizing line sampling for antialiasing.

Contacts: Ray Jones, Ron Perry June 2, 2000
Reference: http://www.merl.com/reports/TR2000-21/index.html

Shadermaps



See Color Plate #2

A frame from the Shadermap enhanced renderer

Most current 3D rendering systems are very inefficient in their operation. Rendered scenes seldom change much between frames, but rendering engines treat each frame as if it were entirely different from the previous one, rendering it from scratch. Shadermaps takes advantage of frame-to-frame coherence in graphics animations by breaking the shader computation into multiple phases. Dramatic improvements in rendering times can be realized by offloading expensive computations on the early phases, and reusing the results of those computations for many frames afterward. Because the early complex phases can use the system CPU, the latter phases are simple and much more amenable to hardware.

Background and objectives: The number of interactive 3D graphics applications is increasing dramatically, driving an insatiable market for high-quality, high-speed 3D rendering engines. The creative content market, particularly the game market, is forever pushing the speed limits of existing engines. Additionally, the increased use of computer generated imagery in feature films and television is pushing the quality limits of even the high-end engines available today. Consequently, there is a need for high-speed, high-quality rendering engines that is not being fulfilled adequately by the market.

In this project we are developing procedural shading rendering methods for generating very high-quality computer generated images at interactive rates. Systems able to produce animations with these effects have been either unbearably slow (taking minutes or days to render each frame of a feature film) or prohibitively expensive. Shadermaps aim to provide full procedural shading in a consumer-level system that operates at interactive rates.

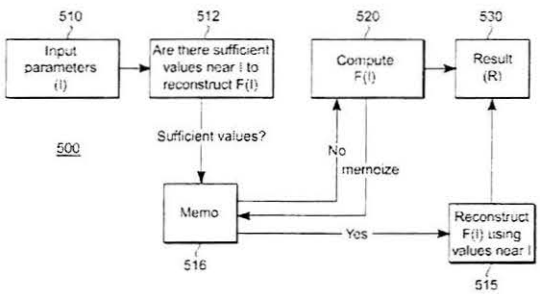
Technical discussion: Animated scenes are comprised of multiple geometric objects with complex appearance provided by procedural shaders. Shadermaps factor each shading computation into static and dynamic phases. The static phase produces an intermediate representation which is processed by the dynamic phase to produce a frame. Due to a clever factorization, the output of the static phase is valid for many frames after its generation, allowing the static phase to run infrequently or even only once for some objects. Since the dynamic phase must run for all objects in each frame, the intermediate representation is chosen so that it can be processed efficiently. To provide interactive frame rates, the intermediate representation and the dynamic phase are designed to be implemented in hardware.

Future Directions: We have implemented a Shadermap enhanced rendering engine in order to explore the data structures suitable for rendering complex scenes with procedural shading. Numerous approaches for factorization are being explored. New representations for procedural shading are also being investigated.

Contacts: Ray Jones, Ron Perry June 2, 2000

Reference: http://www.merl.com/reports/TR2000-25/index.html

Continuous Memoization



Continuous Memoization Flowchart

The term "memoization" is known in the field of computer science as a technique for recording and remembering the results of previous computations so that repeating the same computations can be avoided. The problem with current memoization methods is that the input must exactly match what is contained in the memo, that is the memo is limited to storing and producing discrete values. We define *continuous* memoization as a method that memoizes results of full or partial computations on discrete sets of input parameters. We use the results to reconstruct the computation of an unknown set of input parameters by applying a continuous function on the memoized results. This method avoids recomputing the result for every discrete set of input parameters.

Background and objectives: The new method applies to numerous fields in computer science such as computer graphics. The goal of this effort is to identify which applications can benefit most from the new approach and then to apply the method to those identified domains.

Technical discussion: The new method memoizes a computation as follows. A set of input parameters are provided to the computation. A determination is made to see whether a memo contains results of the computation on sets of memoized parameters near the set of input parameters. If true, the computation for the set of input parameters is reconstructed using the results of the computations on the sets of memoized parameters near the set of input parameters. If false, the computation is performed using the set of input parameters. A result of the computation on the set of input parameters is then memoized, and that result is provided as output of the computation.

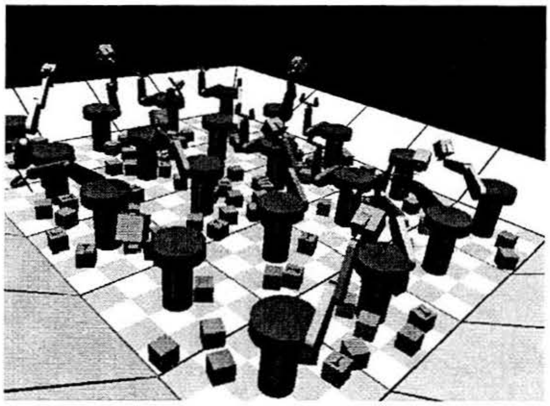
Reconstruction can be done be applying some continuous function, such as interpolation, on the memoized results. Additionally, the computation can be partitioned into a number of partial computations, and only selected partial computations are memoized.

The memo is stored in a cache like memory device. Least recently used replacement (LRU) algorithms can be used to keep the cache at a reasonable size. The cache can be partitioned into tiles, where a tile is the minimal amount of a memo that can be accessed for reading or writing. The size of the tile can be adjusted for a particular computation. In addition, the memo can be maintained at multiple levels of resolution, for example, the tiles can be arranged as a pyramid with coarse, middle, and fine resolutions. Continuous functions can than be applied to derive results for intermediate resolutions, as required.

Future Directions: We are currently exploring all aspects of the continuous memoization approach, such as different reconstruction functions, various cache schemes, and which applications would benefit most from the new technique.

Contacts: Ray Jones, Ron Perry June 2, 2000

Timewarp Rigid Body Simulation



See Color Plate #3

Timewarp Rigid Body Simulation is a paradigm for simulating systems comprising hundreds or thousands of rigid bodies in a physically accurate and efficient manner. The key idea is to evolve the dynamic states as asynchronously as possible, synchronizing states only at the times and to the extents necessary to compute interactions between bodies. Timewarp rigid body simulation facilitates modeling systems that are too large for traditional algorithms, such as factory floors or large, interactive virtual environments.

Background and objectives: Rigid body simulation is difficult because of the discontinuities, such as collisions or contact changes, that affect the otherwise smooth trajectories of state variables. To process discontinuities, traditional high-level simulation loops impose a synchronization among all bodies that is inefficient. The inefficiency is due either to small integration steps or to work wasted when bodies are backed up to the time of a discontinuity. While the inefficiencies are tolerable with small numbers of bodies, the traditional methods become intractable with hundreds or thousands of bodies.

Technical discussion: The timewarp algorithm (Jefferson, 1985) is a paradigm for distributed discrete event simulation based on simulating individual processes independently and judiciously *rolling back* computations when causality violations are detected. With some modification, this approach can be adapted to the continuous-time rigid body simulation problem. Each rigid body or articulated body is a separate process, and these processes communicate via collision impulses and contact forces. A key extension of the classic algorithm facilitates the creation and destruction of processes during the course of the simulation; these events occur as contacts form and break. Another rigid body-specific extension is a method to perform broad-phase collision detection between bodies without synchronizing them. This research also addresses implications for controllers, data capture systems, and other programs that interface with the simulation. Special care is needed since advancement of the bodies is not synchronous or even monotonic.

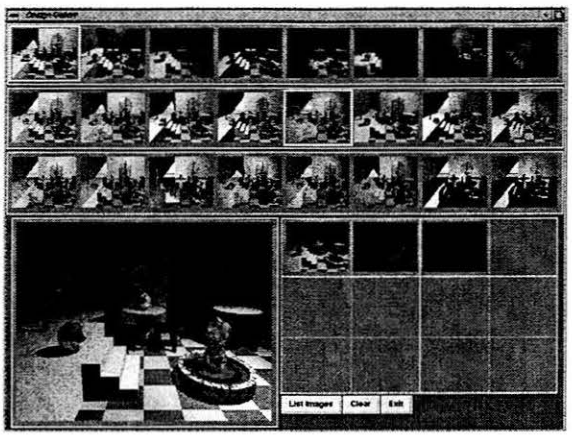
Results from several experiments show that a timewarp rigid body simulator can handle systems with as many as 800 bodies without taking prohibitively small integration steps or wasting excessive work due to rollback. It clearly outperforms traditional algorithms for such large systems.

Future Directions: Currently, timewarp rigid body simulation has only been explored in a uniprocessor setting; more gains could be realized in a multiprocessor or distributed setting.

Contact: Brian Mirtich — March 16, 2000

Reference: http://www.merl.com/projects/rigidBodySim

A Design-Gallery Approach to Lighting Design



DG: Design Gallery *See Color Plate #4*

Good lighting is crucial for high-quality graphics. However, it is still not possible to manipulate virtual lights interactively and see their true effects in real time. Thus the selection and placement of virtual lights is usually a tedious trial-and-error process.

In the DG approach the selection and placement of lights is done jointly by the computer and user. The DG interface presents the user with the broadest selection---automatically generated and organized---of lights that might illuminate the scene. The user can then select individual lights from this gallery and combine their effects in real time.

Background and objectives: The trick in producing compelling lighting designs is to find lighting parameters that yield desirable output values. Tweaking lighting parameters to this end is a tedious experience familiar to anyone who has created computer graphics.

Inverse design is a general paradigm for computer-aided design of graphics. Instead of tweaking parameters, the user supplies an objective function over the output values. This function generates a high score (say) for good output values, and a low score for bad output values. The computer then searches for a set of parameters that will maximize the objective-function score. This general approach has been applied to many problems in computer graphics, including lighting design, but with only limited success. A primary reason for failure is the difficulty of specifying a suitable objective function. In general it is very hard to state mathematically the desirable properties of a lighting design.

Therefore instead of asking the computer "What's best?" we ask the computer "What's possible?" The computer's task is to pick a set of lighting parameters that spans the space of output values (images) as much as possible; the user's task is simply to select from among the presented possibilities. This is the essence of the DG approach.

Technical discussion: Technical details are described in a 1997 SIGGRAPH conference paper.

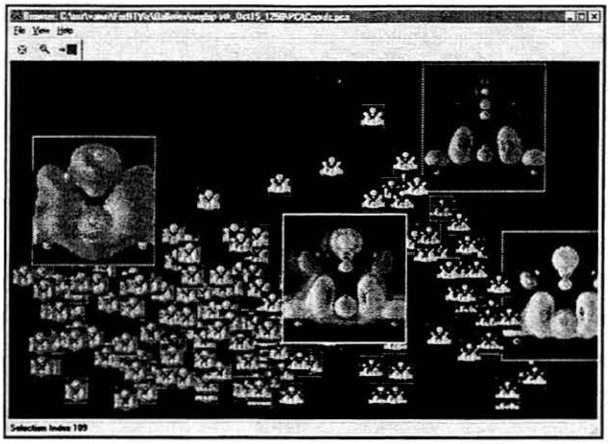
Collaboration: Initial research on the DG/Lighting concept involved researchers from MERL Cambridge Research, Harvard University, and the Georgia Institute of Technology.

Current Status: We have developed a DG plug-in to support lighting design in 3D StudioMax, the most popular 3D modeling and rendering software for PCs, with 86,000 copies sold. This plug-in is being sold through Digimation (www.digimation.com), a software-publishing house that specializes in 3D StudioMax plug-in software. The price is $99.95. Comparable plug-ins have sold between 300 and 2,000 copies.

Contact: Joe Marks June 29, 2000

Reference: http://www.merl.com/projects/design/index.html

VolDG – Design Galleries for Volume Graphics



See Color Plate #5

Direct volume rendering is a key technology for the visualization of large 3D datasets from scientific or medical applications. Of particular importance to the quality of direct volume-rendered images are transfer functions. A transfer function assigns optical properties, such as color and opacity, to original values of the dataset being visualized. Unfortunately, exploring different transfer functions is a tedious task, often done by trial and error. Managing and organizing the exploration of transfer-function space is usually left to the user; the computer is used as a passive instrument. In the Design Gallery for Volume Graphics (***VolDG***) approach the parameter-setting task is divided more evenly between user and computer. ***VolDG*** interfaces present the user with the broadest selection – automatically generated and organized – of perceptually different images that can be produced by varying transfer functions. Our current system is built on top of the popular Visualization Toolkit (***vtk***) and Mitsubishi Real Time Visualization's ***VolumePro 500*** ™ board. The real-time volume-rendering speed of the ***VolumePro 500*** ™ board allows large galleries to be generated in minutes.

Background and Objectives: The trick in producing compelling and useful visualizations is to find input parameters that yield desirable output values. Tweaking input parameters to this end is a tedious experience familiar to anyone who has created computer graphics and scientific visualizations. ***VolDG*** presents a viable alternative to facilitate parameter selection. Instead of asking the computer "What's best?" we ask the computer "What's possible?" The computer's task is to pick a set of input-parameter vectors that spans the space of output values as much as possible; the user's task is simply to select from among the presented possibilities.

Technical Discussion: The principal technical challenges posed by the ***VolDG*** approach are dispersion (finding a set of input-parameter vectors that optimally generates very dissimilar output values) and arrangement (arranging the resulting designs for easy browsing by the user). For dispersion we use a form of evolutionary computation. For arrangement we use multidimensional scaling. Details of our technical approach are described in a 1997 SIGGRAPH conference paper available at http://www.merl.com/projects/dg/. The dispersion process can require the rendering of hundreds or thousands of candidate images, and therefore benefits greatly from hardware acceleration by ***VolumePro 500*** ™.

Collaboration: Joe Marks, Mississippi State University, The Ohio State University, and University of Virginia.

Future Directions: We will make a freeware version of VolDG available on the internet. Based on user feedback, we will improve the interface and widen the application areas.

Contact: Hanspeter Pfister June 1, 2000

Human-Computer Collaborative Optimization

We have developed novel interaction mechanisms that allow people and computers to cooperate in order to solve optimization tasks. We use a tabletop display that allows multiple users to collaborate with each other as well as with the computer.

There are two advantages of our approach. First, the human users can guide systems to solutions that satisfy various real-world constraints. Second, human users can use their visual skills, ability to learn, and strategic sense to improve the performance of the computer search algorithm.

Background and objectives: The motivation of this project is to combine the skills of people and computers to better solve transportation and logistics problems.

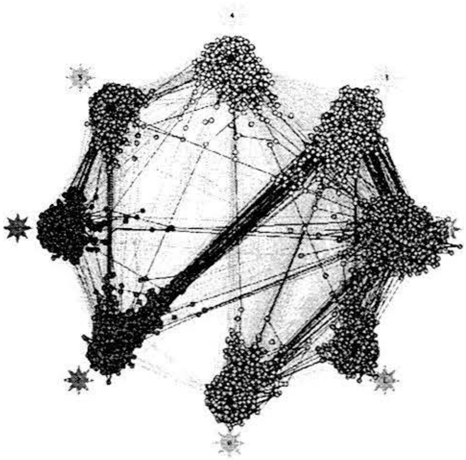
Technical discussion: Typical algorithms for solving optimization problems combine some form of gradient descent to find local minima with some strategy for escaping nonoptimal local minima. and traversing the search space. Our idea is to divide these two subtasks cleanly between human and computer: in our paradigm of Human-Guided Simple Search (HuGSS) the computer is responsible only for finding local minima using a simple hill-climbing search; using visualization and interaction techniques, the human user identifies promising regions of the search space for the computer to explore, and intervenes to help it escape nonoptimal local minima. We have applied our approach to the problem of capacitated vehicle routing with time windows, a commercially important problem with a rich research history. Despite its simplicity, our prototype system is competitive with the majority of previously reported systems on benchmark academic problems, and has the advantage of keeping a human tightly in the loop to handle the complexities of real-world applications.

Collaboration: We are collaborating with the Power and Plant Systems department in Sanken and with Simon Fraser University.

Future Directions: We are currently developing an interactive optimization system to solve graph partitioning problems (which can be applied to chip and board layout). We are expanding our vehicle routing system to handle larger problems and real-world constraints. We are looking for other optimization problems within MELCO.

Contacts: David Anderson, Neal Lesh, Brian Mirtich, Joe Marks February 29, 2000

Human-Computer Collaborative Graph Partitioning



We have developed novel interaction mechanisms that allow people and computers to cooperate in order to solve optimization tasks. We use a tabletop display that allows multiple users to collaborate with each other as well as with the computer.

We first applied our approach to the problem of capacitated vehicle routing. We are now applying these techniques to the problem of graph partitioning, which has application to chip and board layout. We are currently testing our techniques on graphs with 50,000 nodes.

Background and objectives: The motivation of this project is to combine the skills of people and computers to better solve graph partitioning problems, with application to chip and board layout.

Technical discussion: Typical algorithms for solving optimization problems combine some form of gradient descent to find local minima with some strategy for escaping nonoptimal local minima. and traversing the search space. Our idea is to divide these two subtasks cleanly between human and computer: in our paradigm of Human-Guided Simple Search (HuGSS) the computer is responsible only for finding local minima; using visualization and interaction techniques, the human user identifies promising regions of the search space for the computer to explore, and intervenes to help it escape nonoptimal local minima.

The problem of graph partitioning is, roughly, to divide a (typically large) number of nodes into a (typically small) number of partitions, minimizing the number of edges that span two partitions. Our visualization (shown above) allows the human users to view graphs that contain tens of thousands of nodes. Additionally, the user can focus in on on any two partitions to better view the detailed relationships between the nodes in the selected partitions. In our system, the user can perform three operations on the current solution during the interactive search. First, the user can manually edit the current solution by moving nodes or groups of nodes from one partition to another. Second, the user can launch one of many standard refinement algorithms on a focused subset of nodes and partitions. Third, the user can backtrack to a previous solution.

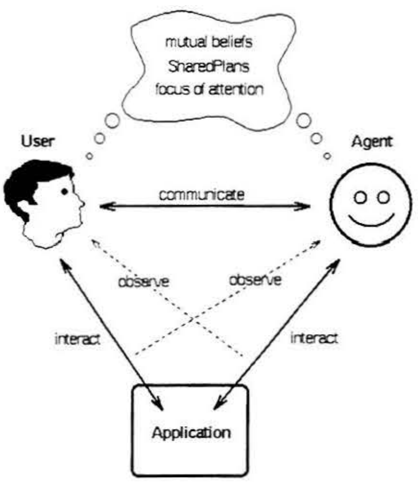
Our preliminary experiments show that users can identify pairs of partitions to focus a computer search on that are, on average, more promising than randomly chosen pairs.

Collaboration: We are collaborating with Simon Fraser University.

Future Directions: We are still developing this application, and our future plans involve implementing several improvements to the current visualization and underlying algorithms. We plan to evaluate our system, and the impact of human guidance on the quality of the final partition, on graphs containing 100,000 or more nodes.

Contacts: Neal Lesh, Joe Marks May 16, 2000

COLLAGEN: Java Middleware for Collaborative Interface Agents



COLLAGEN™: COLLaborative AGENts.

Computer users want the software they interact with to be more than just a set of tools. They want the software to take an active role in their tasks: to advise them when they get stuck, to suggest what to do next when they are lost, and to take care of low-level details after they make a high-level decision. The collaborative interface agent paradigm illustrated at the left achieves this goal by adding a software agent to existing graphical user interfaces. Notice that the user and the software agent can both communicate with each other and observe each other's actions. COLLAGEN is Java middleware for developing such agents.

Background and objectives: COLLAGEN is based on the study of naturally occurring human collaboration, such as two people assembling a complex mechanical device or two computer users working on a spreadsheet together. Our objective is to provide collaborative agents for the training and/or use of any complex software system.

Technical discussion: COLLAGEN automatically maintains a history (log) of the user's and agent's activities, which is hierarchically segmented according to the structure of the tasks they are performing. This structure is used to help users when they get confused and to remind them what they was doing after an extended absence.

To apply COLLAGEN to a particular software system, the system developer must provide an abstract model of the kinds of tasks for which the system will be used. This knowledge is then automatically compiled for use by the interface agent.

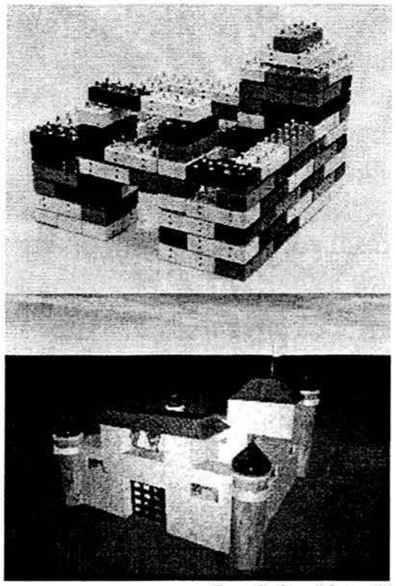
Collaboration: We are currently working with the Systems Technology Dept. of IESL (Sanken) to build a COLLAGEN-based agent to help users of their GhostHouse software tools. We are also working with the Univ. of S. California, MITRE Corp., and U. Pittsburgh on tutoring applications and with Delft U. of Technology on applications to intelligent consumer products.

Future Directions: We hope to demonstrate the power of COLLAGEN by implementing agents for several different kinds of applications in the next year or so. Some of these applications will require solving new basic research problems in collaborative discourse. In addition, we are beginning to explore how to use COLLAGEN together with off-the-shelf speech recognition and generation technology to build conversational speech systems.

Contacts: Charles Rich, Neal Lesh, Candace Sidner April 15, 2000

Reference: http://www.merl.com/projects/collagen

Architectural Interpretation of 3D Models



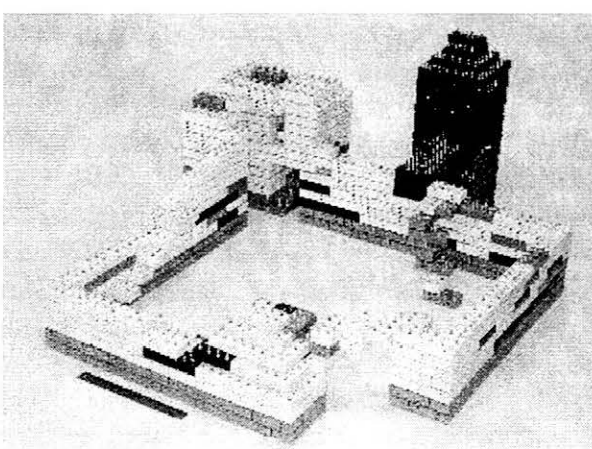
See Color Plate #6

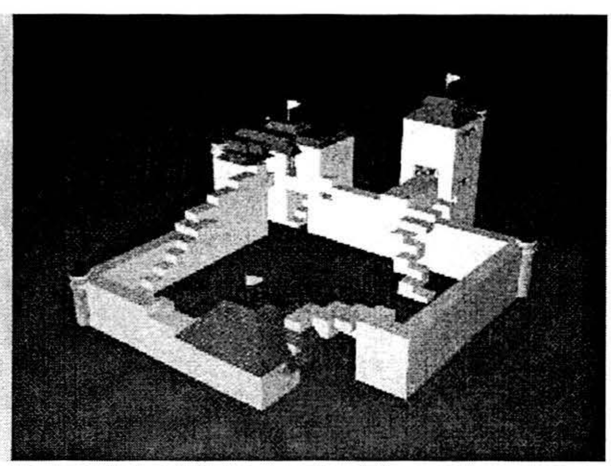
Using our self-describing building blocks, anyone can create 3D geometry, but creating fully realized, detailed models remains a difficult and tedious task. Our system shows how a computer might assist by automatically identifying architectural elements in a 3D model of a building, and rendering those elements in various styles, with decorative interpretations.

Above at left is a physical model made from 118 blocks. Each block contains a PIC microcontroller and communicates with its neighboring blocks; a connected PC computes the geometry from connectivity information uploaded from the block structure. Below at left the same building is shown, having been automatically interpreted and rendered in a castle style.

Background and objectives: Possible applications for this technology range from creating virtual environments for game playing to architectural studies.

Technical discussion: A rule-based system implemented in Prolog recognizes the major architectural elements in structures made from the blocks, such as the roof, walls, and corners, and then assigns materials and adds decorations in keeping with the style or theme selected by the user. Below is another example of the same castle style applied to a structure built from 560 blocks.





Collaboration: The University of Virginia worked with us to develop the decorative renderings.

Contacts: Joe Marks, Kathy Ryall, David Anderson, Jamie Frankel. June 2, 2000

Reference: http://www.merl.com/reports/TR2000-13/index.html

Learning Concise and Minimally Uncertain Models

$$e^{-H(\theta)}$$

Entropic estimation is a framework for simultaneously estimating the structure and parameters of a probabilistic model. It is a continuous formalization of Occam's razor: Seek the smallest and most unambiguous model that can explain the data. The resulting models are more predictive, more general, and much more interpretable than models obtained from conventional learning methods. Entropic estimation often induces a model quite close to the mechanism that generated the signal.

Background and objectives: To tune in a radio station, you get a machine—a radio—and you vary a parameter—its frequency—until the machine has a good fit with the signal you want. Computer "learning" is very similar: The computer gets a machine—usually a complex statistical model—and programmatically varies hundreds or thousands of parameters until the model fits the signal. There is a catch: An expert must design a model that matches the gross structure of the signal. This means laborious trial-and-error, sometimes without success. We seek efficient methods to simultaneously design the machine and find good parameter settings.

Technical discussion: Entropy is a measure of disorder, spread, and uncertainty. We learn by minimizing three entropies, one assessed on the model, one assessed on the model's picture of the data, and one assessed on aspects of the data not captured by the model. There is an intuitive probabilistic interpretation in terms of maximizing the posterior given by Bayes' rule, which measures our confidence in a model after having seen some data:

$$P(\textit{model given data}) = P(\textit{data given model}) \bullet P(\textit{model}) \,/\, P(\textit{data})$$

The likelihood *P(data given model)* is a known function $f(X, \theta)$ of data X parameters θ. The prior *P(model)* is usually unknown. Entropic estimation gives us a universally applicable prior,

$$P(\textit{model}) \propto \exp(-H(\textit{model}))$$

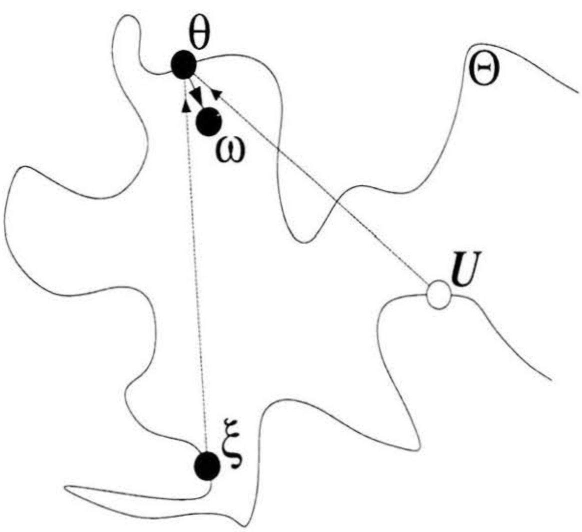
where *H()* measures the entropy, or uncertainty, of the model. The prior summarizes our background knowledge--this particular prior can be derived from the simple assertion, "This task is learnable." It is also a bias for small, unambiguous, and highly structured models.

An estimator yields parameter values that maximize the posterior. We have derived exact solutions for estimators for a large variety of likelihood functions--even when the problem leads to systems of transcendental equations. These yield very fast learning algorithms that sculpt theories out of overcomplete random models by extinguishing excess and inappropriate parameters, thereby removing terms from the likelihood function $f(X, \theta)$. Variants on the framework give trimming criteria that tell us when the model can be simplified by removing parts of the likelihood function, and deterministic annealing procedures that allow us to avoid suboptimal solutions. Because entropy minimization will find hidden structures in the data, it can be thought of as automated exploratory science, discovering previously unknown hidden causes that explain our observed world.

Contact: Matthew Brand
http://www.merl.com/projects/entropic/

May 18, 1999

Cross-entropy Estimation



Statistical models lie at the heart of advanced technologies such as communications, sensing, language processing, and signal processing; they are important to any industrial process that depends on data collection and analysis. The better the model, the better the process that depends on it. We present a basic result on the optimal bias for learning these models from data, and practical algorithms that translate this result into superior statistical models. Immediate applications include better classifiers (e.g., for predicting disease from weakly diagnostic tests), and a method for manipulating stylistic qualities of data-sets such as those used for character animation.

Background and objectives: When forming a statistical model from data, a density model generally contains the most information and can support the broadest variety of queries. A density model describes how the data-points are distributed through the space of possible measurements, including where they are dense. However, specialized models often outperform density models for limited kinds of queries such as classification, largely because it is much harder to estimate a good density model than a classifier.

Technical discussion: Density estimation is a kind of inference; every inference begins with a prior belief. We showed that the optimal prior belief for extracting information from data is a preference for the least uniform density that describes the distribution of the data well. Formally, one wants to minimize the entropy (uncertainty) of one's model of the data, or, equivalently, maximize its cross-entropy with the uniform (zero-information) density. In this light, we developed efficient algorithms for estimating a density $\boldsymbol{\theta}$ from data $\boldsymbol{\omega}$ while maximizing cross-entropy (a measure of difference) with another density ξ (and/or the uniform density $\boldsymbol{U}$). This is depicted in the figure above where $\boldsymbol{\Theta}$ represents the manifold (surface) of all possible models available. These techniques have numerous applications, including estimating models that identify stylistic variations in datasets (see the "Style Machines" URL below), and estimating density models whose classification accuracy is competitive with the best current specialized classification methods. By maximizing the cross-entropy between density models for each class of data, we minimize their overlap and the probability of incorrectly classifying any points near their boundary.

Contacts: Matthew Brand June 1, 2000

Reference: http://www.merl.com/reports/TR2000-14

Voice Puppetry

See Color Plate #7

The voice puppet can animate any face using just your voice. It uses expressive information in a voice-track to control the entire face, from lips to eyebrows, neck to hairline. The mapping from vocal to facial gestures is learned from vision of real facial behavior and automatically incorporates long-term vocal and facial dynamics such as co-articulation. The animated face can be a 2D cartoon, a 3D model, or even a photo.

Voice puppetry is intended to replace tedious and expensive methods currently used in cartoon animation, film special effects, and video post-production. In addition, it should create new opportunities for realistic facial motion in video games and computer entertainments.

Background and objectives: Nearly all facial animation systems begin with a stream of phonemes (basic sound tokens), usually obtained by hand, from text, or, less successfully, from speech recognition. Typically, each phoneme is mapped to a viseme (facial pose) which are interpolated to produce an animation. It is widely understood that phonemes and visemes are inadequate because they discard information about expression and emotion in the whole-face gesture. Even for the limited problem of lip-syncing, phonemes and interpolation lead to problems with unnatural facial dynamics.

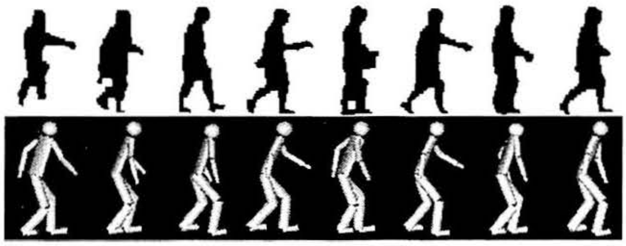
Technical discussion: We learn the natural dynamics of the whole face by modeling the motions of facial with an entropically estimated hidden Markov model. Entropic estimation produces a compact, sparse, and minimally ambiguous state machine, essentially discovering key facial states, dynamics, and timing. Given a new vocal track, the system calculates a trajectory through facial configuration space that is maximally compatible with the learned facial dynamics and with the newly observed acoustic features. The resulting trajectory can be used to drive a variety of animations, ranging from 2D cartoons to 3D computer graphics to 2.5D image warps. Our current system animates a texture-mapped 3D model, producing a surprisingly good illusion of realism. We have also used the learned model to animate non-human heads and for extremely low bit-rate facial motion coding (as low as 4 bits/frame!).

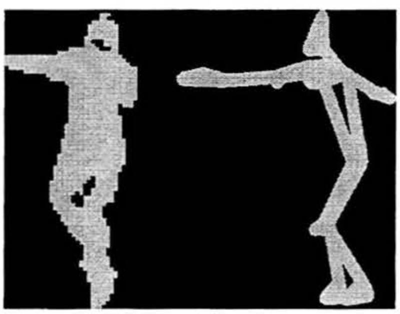
Future Directions: MELCO's Digital Media Lab is evaluating voice puppetry for the Japanese entertainment industry. American animation studios are also seeking licenses.

Contact: Matthew Brand

Reference: http://www.merl.com/projects/puppets/ June, 2000

Shadow Puppetry





The shadow puppet infers full 3D body pose and orientation from a sequence of silhouettes. It provides a low-cost form of motion capture that could be used for video games, film special effects, and surveillance. Simple and robust computer vision algorithms compute a sequence of silhouettes from video input. The shadow puppet infers a 3D motion sequence that is most consistent with the silhouettes and with prior knowledge about how the body moves. Everything is learned from data. If the puppet is trained on stylistic motion, it will infer motion in that style, e.g., if trained on a good dancer, when observing a poor dancer it will synthesize more graceful motion.

Background and objectives: Motion capture data is the basis for a wide variety of film and video special effects, as well as for synthetic characters if videogames and virtual spaces. It is also extremely expensive: Full-body motion capture systems cost $50K and up, and often produce "dirty" data that requires days of cleaning by highly skilled technicians. Our goal is to produce high-quality data using a PC with low-cost consumer-quality video input.

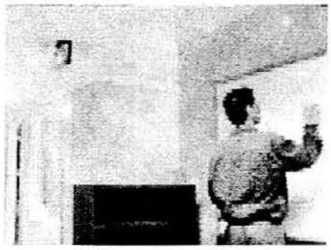
Technical discussion: We learn the natural dynamics of the whole body by modeling existing motion capture data with an entropically estimated hidden Markov model. Entropic estimation produces a compact, sparse, and minimally ambiguous state machine, essentially discovering key body poses, dynamics, and timing. This model captures the kinematics, dynamics, and style of the original performers. The model is made to observe both the 3D data and simple statistics of its 2D silhouettes. These silhouettes can be reliably extracted from video if the background is stable. Given new video data, the system calculates a trajectory through 3D configuration space that is maximally compatible with the learned body dynamics and with the newly observed silhouettes, essentially synthesizing motion capture data.

Future Directions: We are looking at combining this technology with stereo vision.

Contact: Matthew Brand

Reference: http://www.merl.com/projects/shadows/ May 18, 1999

Learning Normal Activity and Detecting Anomalies



writing

on the phone

At present, you attend to the computer. In the future, the computer should attend to you. This computer vision system watches you and learns your normal work habits and basic activities. It can summarize your day, adjust your environment for different of activity, or signal an alarm when something unusual or disruptive happens.

Abbreviations: *HMM: hidden Markov model; MAP: maximum a posteriori.*

Background and objectives: We are interested in the use of computer vision to "keep an eye on things" while you're busy or away. This application, a personal security device, might be used to keep an eye on a child at play, a hospital patient, or an elderly parent.

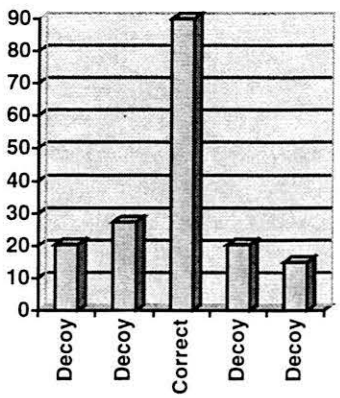
Technical discussion: The basic approach is to learn a probabilistic model of the scene dynamics. The demonstration system extracts a motion-based feature vector from each video frame, then learns a summary of signal dynamics in the form of a hidden Markov model. An entropy-minimization algorithm estimates both model structure and parameters, yielding a small, sparse, accurate model that generalizes very well to new video. Due to its sparsity, the states in the model are highly correlated with meaningful partitions of the video, e.g., activities. The resulting model reads like a flowchart of a normal day's work activities. We can use the model to detect when you are in particular states, e.g., working at the whiteboard, and have the computer adjust your environment appropriately, e.g., brightening the lights. The system is also very good at detecting anomalous activity—the laboratory prototype proved quite adept at detecting coffee buzz an hour after several espressos.

Future Directions: This approach has also been applied to monitoring traffic at intersections and is now under consideration for "smart buildings" that monitor activity in their public spaces.

Contacts: Matthew Brand February 7, 2000

Reference: http://www.merl.com/projects/activity

Protein Recognition



The protein folding problem (predicting the three-dimensional configuration of a protein given the sequence of amino acids) is one of the "holy grails" of computational biology. Experimental techniques to determine protein structures such as x-ray crystallography are not able to keep up with recent huge improvements in our ability to sequence DNA (and thereby proteins), making a theoretical solution to the problem even more urgent. A solution to the problem would have extraordinary scientific implications and open the way to rational drug design.

An important sub-problem that is easier to analyze is the "protein recognition" problem—recognizing the correct structure of a protein from many plausible "decoy" structures. If one can construct a scoring function which scores the correct structure much higher than very plausible decoys, one can use that scoring function to solve the protein folding problem.

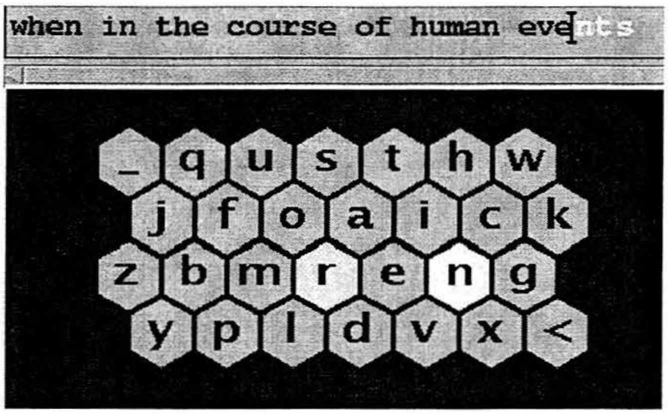
Background and objectives: The discriminative power of a protein recognition model is usually specified by its "z-score", which measures how many standard deviations above the mean score of the decoy structures it scores the correct structure. The z-score will depend on the database of decoys used, but we were able to develop a scoring function that typically improved the z-score by 10-20% (e.g. from a z-score of 6 to one of 7) compared to the best protein recognition scoring functions that had previously been developed.

Technical discussion: The novel idea underlying this work was to broaden the space of energy functions for which one could rapidly and accurately estimate the z-score. In particular, an energy function was used which combines a standard two-body energy term with a new term that depends on the secondary structure of the protein. Because the z-score could be estimated rapidly, one could then search over the space of all such models for the one that gave the best z-score.

Collaboration and Future Directions: The work so far has been done internally at MERL Cambridge Research, but for the future, it would make sense to collaborate with the theoretical protein folding group at Harvard University's Chemistry department, with whom we have had numerous useful discussions.

Contact: Jonathan Yedidia (with advice from Matt Brand) March, 2000

The Tapwriter



The Tapwriter keyboard is a text entry interface that is optimized for fast typing with a single pointer such as a finger, pen, computer mouse, or eye gaze. It is suitable for very small displays such as seen on cellphones, palmtops, and wearable computers (e.g., a form-factor like a wristwatch), as well as for handicapped-access devices.

Background and objectives: With the rise of internet connectivity, it will be necessary to provide some means of text entry for small computing platforms such as cell-phones. Our goal is to design a text-entry interface that maximizes the speed of hunt-and-peck typing but minimizes real estate.

Technical discussion: English text contains roughly 1 bit of information per character; if asked to predict the next character in a text, you can use the context (spelling, grammar, meaning) to whittle down the alphabet to an average of 2 or 3 possible completions. The Tapwriter uses machine learning methods to reduce typing to a choice between two or three alternatives. These are highlighted, as shown in the picture ("n" and "r" are highlighted after the user has written "eve"). The user can then make a choice with a pointer or a small number of buttons.

The Tapwriter is designed to minimize the work that the eye does in searching for the right key and the travel that the hand does in selecting it. The layout of the keyboard is optimized to minimize the average distance traveled while typing English text. Most of the time, the letter the user intends to type next adjoins the letter just chosen. On average, the pointer will travel 1.6 key-widths per letter while typing standard English. (Travel on QWERTY keyboards averages 4.5 key-widths per letter.)

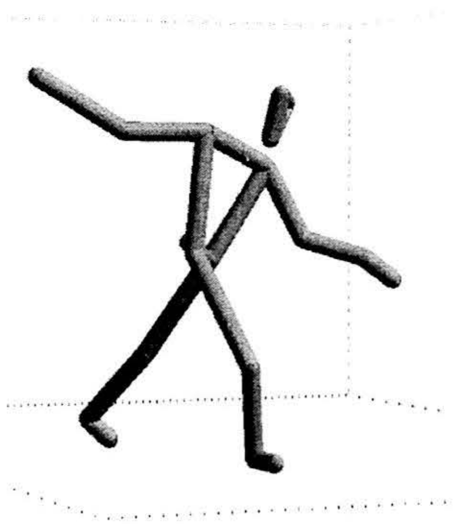
Because the layout is unfamiliar, variable length Markov models are used to predict what the user will type next. The choices are highlighted to lead the eye to them, to reduce hunting time. The predictions are correct 76% of the time for normal English text. Novice users can write 40-80 characters per minute with the Tapwriter, as fast as expert typists can write using the familiar layout of the QWERTY keyboard on a pen-based device such as a PalmPilot. With practice one can tap-type faster than 140 characters/minute, roughly half the speed of touch-typing. The Tapwriter can also suggest completions of words (shown in white to the right of the cursor), based on the user's habitual language.

Future Directions: A software prototype was developed in January 2000. ID-ken and Johosoken are evaluating Tapwriter technologies for cell-phone and handwriting recognition products.

Authors: Matthew Brand
Reference: http://www.merl.com/people/brand/

February 1, 2000

Synthesizing Stylized Motion Data for Animation



Style machines are statistical models of human motion that can synthesize high-quality motion data for animation in video-games, movies, and simulations. Their outstanding feature is that they can synthesize the same qualitative motion in a broad variety of styles, giving the animator control over a performance's mood, energy, grace, and even masculinity/femininity. They can resynthesize existing motion-captured data in new styles, or synthesize wholly new performances, eliminating costly motion-capture studio sessions from production schedules and budgets.

Abbreviations(s): mocap: motion-capture

Background and objectives: The fast growing animation and special-effects industries have a voracious appetite for motion data. The gold standard is motion data captured directly from the body of a trained actor in an instrumented studio. However, mocap is extremely expensive, and the results are often unsatisfactory. Our objective is to synthesize and resynthesize virtual mocap data, giving the animator explicit control over the choreography and style of the output.

Technical discussion: A style machine is a generative model of human 3D motion that is learned from old mocap. The learning algorithm breaks the data down into small motion primitives, stylistic variations in how those primitives are performed, and rules for sequencing those primitives to form plausible choreography. Part of the analysis identifies a small number of stylistic degrees of freedom that account for all the variations in the training set.

An animator then chooses a setting of the style parameters and specifies some choreography to perform. The style machine synthesizes virtual motion capture in that style. With appropriate training, the style machine can extract the choreography automatically from an arbitrary signal, for example, old motion capture or video.

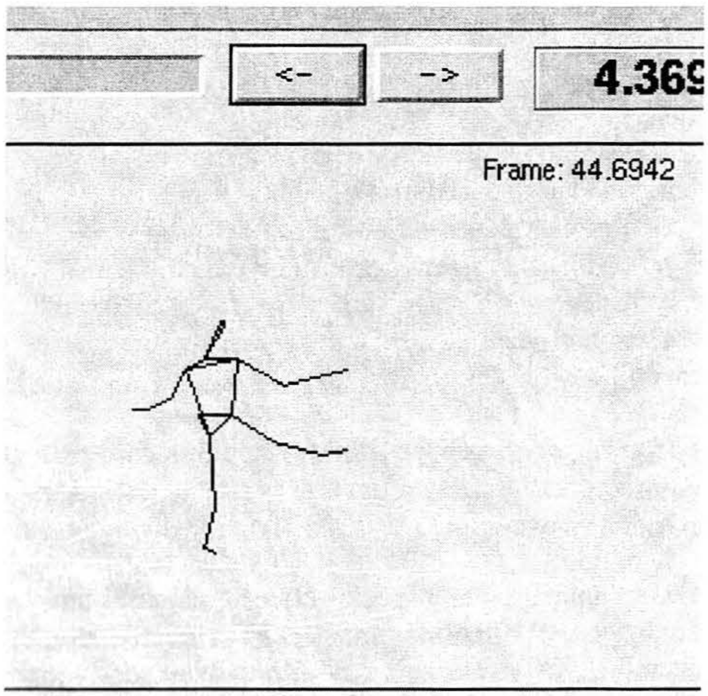
Collaboration: Work done with graphics talent from New York University.

Future Directions: A working prototype was trained in the fall of 1999. We are collecting data for a more extensive trainng session and exploring applications within MELCO's DML subsidiary.

Contacts: Matthew Brand May 12, 2000

Reference: http://www.merl.com/reports/TR2000-14

Synthesizing Human Animation from Examples



Realistic looking human animation is notoriously difficult to generate and control. Here we use a training based approach to synthesize realistic looking animation. We first gather a long sequence of human motion data, from 3-d motion capture technology. Then we break the motion data into snippets of ½ or 1/3 of a second each, and translate and rotate each snippet into a canonical position. To synthesize a new animation, we select snippets from the database of training examples and concatenate them together to form the synthesized output. The selection is based on how well the snippet fits with the previous frame's snippet, and how well it meets some objective function inputted by the user. That objective function might describe the desired body position or configuration of the character. The result is a realistic human animation, conforming to the required objective function as well as possible, given the training data.

Background and objectives: The method falls in the category of data driven synthesis. All of the output motions look realistic because each is taken from the actual motion of an actual dancer or actor. The only processing of the motions that takes place is a translation and rotation in order to place the dance snippet in the appropriate position and orientation for motions constructed so far in the synthesis.

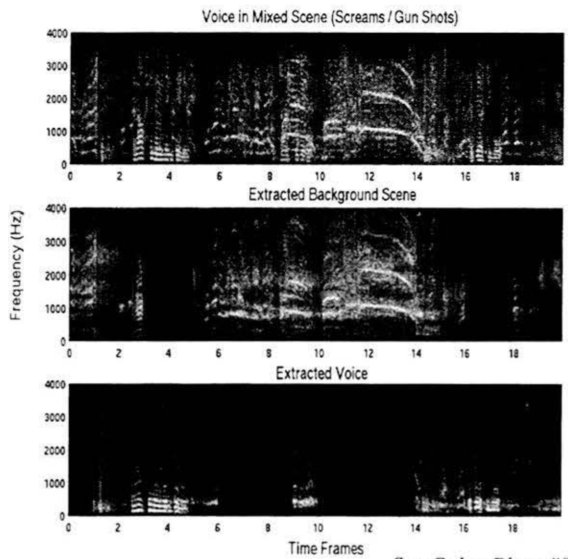
Technical discussion: There is a search involved to determine which snippet to select at each given time step. At the cost of computational complexity, one can look ahead one or two time steps in order to select the snippet for the current time step which will lead to best meeting the objective function several time steps in the future. This modification to the greedy algorithm (looking ahead no time steps) improves the synthesized result considerably.

Collaboration: Psychology Dept., Stanford University.

Future Directions: This approach may be extended to modelling other complex dynamical systems, for example, the motion of a falling piece of paper. With enough samples, one can make the paper fall in a realistic manner, yet at the same time conforming to some objective function describing the desired path of the paper through space.

Contact: Bill Freeman March 16, 2000

Sound Spotter: Recognition and Extraction from Mixed Audio



See Color Plate #8

Sound Spotter is a system for analysis and tracking of individual sound sources in mixed audio scenes. This technology has applications in artificial hearing and audio content description.

Artificial hearing addresses the problem of speech recognition in everyday environments, such as the home or a busy street. Applications include improved speech interfaces and electronic assistance devices for the hearing impaired.

Additionally, the core technology has been included in the MPEG-7 standard for indexing and extraction of audio content in digital media applications.

Abbreviations: ASR Automatic Speech Recognition, ICA Independent Components Analysis

Background and objectives: Audio recognition systems, such as ASR and Internet audio search engines, require that a sound signal is isolated from all other sounds, such as noise in the environment. But sounds in the real world occur in context such as shown in the figure above; a news broadcast is delivered against a background of screams, gunshots and crowd noise. The Sound Spotter isolates and separates independent source sounds from the mixture; in this case the speech and the screams are separated. This technique may enable the use of ASR in non-ideal environments; e.g. a conversations in a busy street or office, or recognition of actors' voices in a film sound track. Identification and extraction of non-speech sounds, such as musical instruments or sound effects, is also enabled using this method.

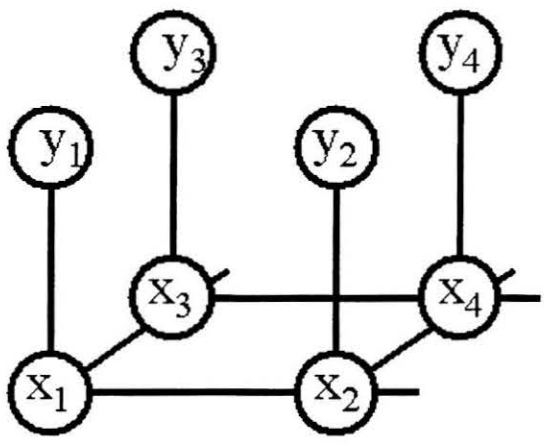
Technical discussion: Sound Spotter employs a higher-order statistics framework, using ICA, in order to decompose the output of an auditory filter-bank into statistically independent features. These features are organized into clusters using mean-field theory annealing. The resulting clusters are groups of sound features that correspond to persistent underlying sources: such as a segment of speech or a gunshot. One major advantage of the technique is that a single mixed audio channel is sufficient for successful operation. This contrasts with blind signal separation algorithms that generally require multi-channel input and are thus limited in application scope.

Collaboration: The Sound Spotter methods have been included in the working draft of the MPEG-7 international standard and development of the MPEG-7 experimental model contribution is a collaborative effort between HSL and MERL.

Future Directions: ASR enhancement, separation from noise, media annotation, artificial hearing.

Contact: Michael A. Casey April 25, 2000

Correctness of Belief Propagation in Bayesian Networks with Loops



Bayesian networks, used in many machine learning applications, represent statistical dependencies of variables by a graph. For example, in the figure, the "y" variables may be observations, and the "x" variables may be unknowns. Local "belief propagation" rules are guaranteed to work properly in networks without loops, but researchers have also found good performance on graphs with loops.

We provide theoretical understanding of this good performance for Gaussian random variables, and for the case of calculating the local maximum of the posterior probability.

These results motivate using the powerful belief propagation algorithm in a broader class of networks.

Background and objectives: Problems involving probabilistic belief propagation arise in many applications, including error correcting codes, speech recognition and medical diagnosis. If the graph is singly connected then there exist efficient local message--passing schemes to calculate the marginal probability of a variable, needed to solve these problems. Pearl (1988) derived such a scheme for singly connected Bayesian networks and showed that this ``belief propagation" algorithm is guaranteed to converge to the correct marginal probabilities (or ``beliefs"). However, the same algorithm will not give the correct beliefs for networks with loops.

Technical discussion: We analyze belief propagation in graphs of arbitrary topology but focus on nodes that describe jointly Gaussian random variables. We give an exact formula that relates the correct marginal posterior probabilities with the ones calculated using loopy belief propagation. We show that if belief propagation converges then it will give the correct posterior means for all graph topologies. We present a relationship between the error in the covariance estimates and the convergence speed. For Gaussian or non-Gaussian variables, we show that the "max-product" algorithm, which calculates the MAP estimate in singly connected networks, only converges to points that are at least local maxima of the posterior probability of loopy networks. This motivates using this powerful algorithm in a broader class of networks.

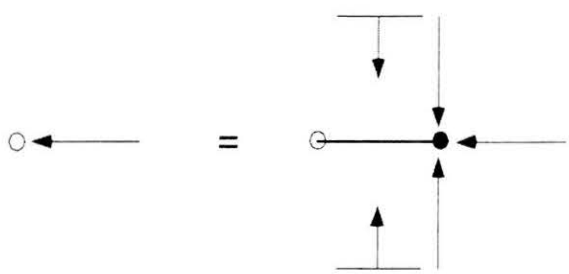
Collaboration: This work was a collaboration with a researcher at the University of California, Berkeley, USA.

Future Directions: We plan to use these results in applications involving computer vision, such as super-resolution and estimating the shape of an object from its image.

Contact: William T. Freeman

Reference: http://www.merl.com/reports/TR99-33 June 22, 2000

Generalized Belief Propagation Algorithms



Many problems in computer vision, machine learning, diagnosis, statistical physics, and combinatorial optimization can be posed in terms of a probabilistic graphical model consisting of a lattice of nodes with links connecting nodes that influence each other. Typically, one asks for the probability that a given node or collection of nodes is in some state, given the states of another set of nodes. Alternatively, one might ask for the most probable configuration of the system as a whole given the state of some of its nodes.

These problems can be solved exactly when the lattice has the topology of a chain or tree, but in most problems of practical interest, the lattice has loops. We have invented a class of message-passing algorithms which can quickly and accurately (albeit approximately) solve these problems for lattices of arbitrary topology. The algorithms have the attractive feature that one can always pay for greatly increased accuracy with more computation, and the computational cost is not excessive.

Background and objectives: We discovered that previously developed message-passing algorithms (in these algorithms, each node sends "messages" to its neighbors about what states they should be in) gave equivalent results to an approximation known in the statistical physics literature as the "Bethe approximation". A class of physics approximations invented by Kikuchi which generalized and improved the Bethe approximation had previously been solved for homogeneous systems like the ferromagnet, but not for inhomogeneous systems such as arise in computer vision. We were able to discover message-passing algorithms that gave results equivalent to the Kikuchi approximations, but which had a form amenable to solution even for inhomogeneous systems. These algorithms therefore represent a break-through in theoretical physics as well as computer science.

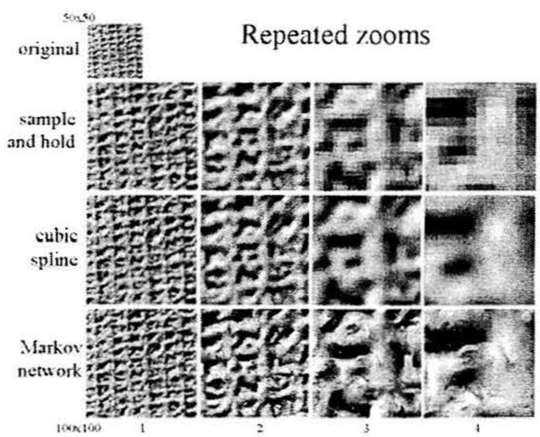
Technical discussion: Intuitively, the new algorithms give more accurate results because they allow clusters of nodes to send messages to other clusters, while always maintaining the appropriate constraints between the probabilities in all clusters of nodes. By choosing larger clusters, one can obtain more accurate results, at the cost of some more computation.

Collaboration: Computer Science at U.C. Berkeley

Future Directions: A patent application is in process. We plan to apply these new algorithms to a host of problems, probably starting with problems in combinatorial optimization like the traveling salesman problem. Essentially any problem which is currently treated by Monte Carlo simulation or simulated annealing could potentially be solved more quickly and accurately by these new algorithms.

Contacts: Bill Freeman and Jonathan Yedidia March 16, 2000

Super-resolution using a Markov network approach



One would like to have an intelligent method for expanding the resolution of an image. It should keep edges, which are implicitly described, in the low-resolution image sharp. It should make intelligent guesses about the details of textures.

We continue the development of our Bayesian method, which addresses those goals. The method estimates the detail components of a scene, given the low resolution components We use a machine learning algorithm to estimate the most probable sharp image that would correspond to the observed low-resolution image.

The focus of our recent work has been on simplifying the training and set-up computations, and reducing artifacts.

Background and objectives: Images are typically represented as collections of pixels, yet we would like to treat them as if they were resolution independent objects. In graphics, we have that option with polygon based representations: if one zooms in on a polygon-defined edge, the edge will stay sharp through all levels of zooming. We would like to have a similar level of resolution independence even in our pixel-based image representations.

Technical discussion: We use a training based approach. We examine many pairs of high resolution, and low resolution versions of the same image data. We divide each image into patches, both high resolution and low-resolution patches. We form a training database of 50,000 – 100,000 high and low-resolution patches.

Given a new low-resolution image, we seek to estimate the most probable corresponding high-resolution image. In the training database, there may be several different examples similar to any given input low-resolution patch. Typically, we gather a collection of 10 candidate high-resolution image patches to explain each low resolution input patch. The requirement of compatibility between the candidates of neighboring patches is used to select which of each patches 10 candidates is the best choice. We have improved the method used to measure compatibility between neighboring high-resolution candidate patches, which speeds up the set-up time needed to process an image.

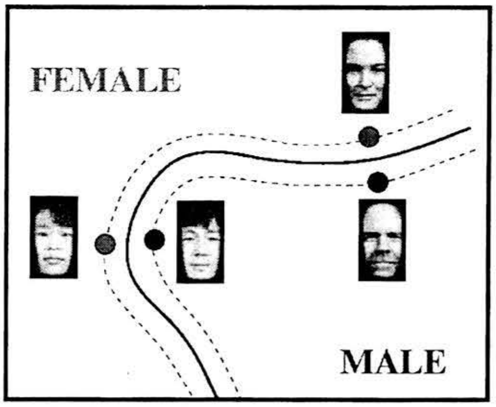
Collaboration: We are collaborating with the Advanced Television Laboratory, in Princeton, New Jersey, USA.

Future Directions: We hope to decrease the time required for processing the images.

Contact: William T. Freeman June 6, 2000

Reference: http://www.merl.com/reports/TR2000-08/

Support Vector Learning for Gender Classification



Computer vision systems for people monitoring will eventually play an important role in our lives by means of automated human (face) detection, body tracking, action (gesture) recognition, person identification and estimation of age and gender. We have developed a facial gender classifier using Support Vector Machine (SVM) learning with performance superior to existing gender classifiers. This technology can, for example, be used for passive surveillance and control in "smart buildings" as well as gender-mediated HCI.

Abbreviations: SVM: Support Vector Machine, HCI: Human Computer Interface

Background and objectives: This project addresses the problem of classifying gender from low-resolution images in which only the main facial regions appear (i.e., without hair information). We wanted to investigate the minimal amount of face information (resolution) required to learn male and female faces by various pattern classifiers. Previous studies on gender classification have relied on high resolution images with hair information and used relatively small datasets for their experiments. In our study, we demonstrate that SVM classifiers are able to learn and classify gender from a large set of hairless low resolution images with the highest accuracy.

Technical discussion: A Support Vector Machine is a learning algorithm for pattern classification and regression. The basic principle behind SVMs is finding the optimal linear (or nonlinear) hyperplane (see above figure) such that the expected classification error for unseen test samples is minimized (i.e., good generalization performance). We investigated the utility of SVMs for visual gender classification with low resolution "thumbnail" faces (21-by-12 pixels) processed from 1,755 images from the FERET face database. The performance of SVMs (3.4% error) is shown to be superior to traditional pattern classifiers (Linear, Quadratic, Fisher Linear Discriminant, Nearest-Neighbor) as well as more modern techniques such as Radial Basis Function (RBF) classifiers and large ensemble-RBF networks.

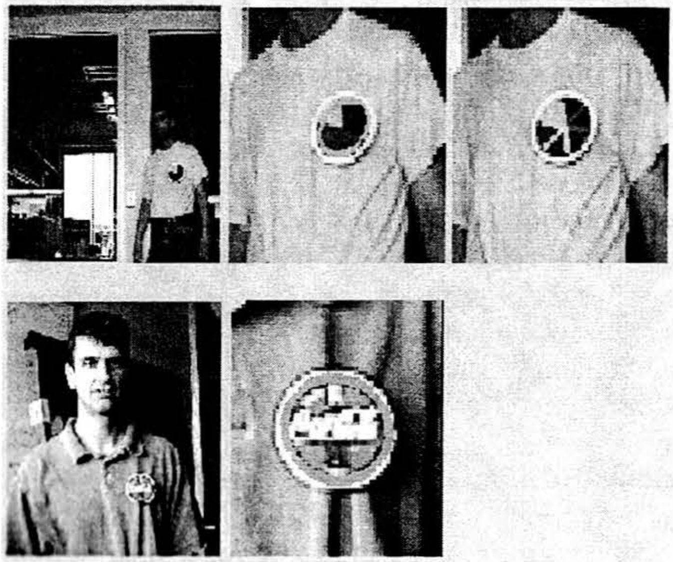
Collaboration: Beckman Institute, University of Illinois at Urbana-Champaign

Future Directions: The SVM-based visual gender classifier developed is not invariant to head orientation and (like many other systems) requires fully frontal facial images. While it functions quite well with frontal views of the face, the inability to recognize gender from different viewpoints is a limitation which can affect its utility in unconstrained imaging environments. However, the system can be of use in application scenarios where frontal views are readily available (e.g., doorways, public kiosks, advertising billboards, etc.).

Contacts: Baback Moghaddam December 15, 1999

Reference: http://www.merl.com/projects/gender/index.html

Detecting Visual Tags



This work addresses the detection by computer vision of visual tags. Examples of visual tags include color-coded tags, which encode identity. It is possible to generate thousands of distinct color codings for a badge of the type illustrated at top-left. Such tags could be worn by people or attached to objects, and are cheap and disposable. Other examples of visual tags are logos. The illustration at bottom-left shows the detection of a Coca-Cola logo. This detection could be used to add intelligence to a Coke vending machine, enabling the machine to issue a special greeting to a customer wearing the logo, or informing the customer of a discount in the price of a soft-drink. A further use of visual tags is to transmit information about the wearer - for example, to indicate that the wearer is infirm, elderly or handicapped, so that a detecting system such as an elevator can offer a special level of service.

Background and objectives: (a) Identity tags: cheap and disposable identity tags have many uses for tagging and tracking objects. (b) Logos: an interactive vending machine, which responds when a customer is wearing a product logo, could increase sales and could encourage customers to wear the logo thereby providing advertising. (c) Information tags: a badge indicating that the wearer is elderly or infirm could be used to automatically invoke special levels of service from the detecting system - for example, an elevator system could level the floor more accurately on arrival if a person in a wheelchair is present.

Technical discussion: The approach is general enough to work for a variety of tag appearances and sizes. The system uses two cues – (a) shape of the tag, and (b) the internal design of the tag. Processing speed is about 1Hz on a standard PC processor, for a 240-x-240 pixel image. The system works for tag sizes in the image down to about 15 pixels diameter. Thus the system can operate to a range of about 10 feet for a 240-x-240 pixel image, and this range increases for higher image resolutions.

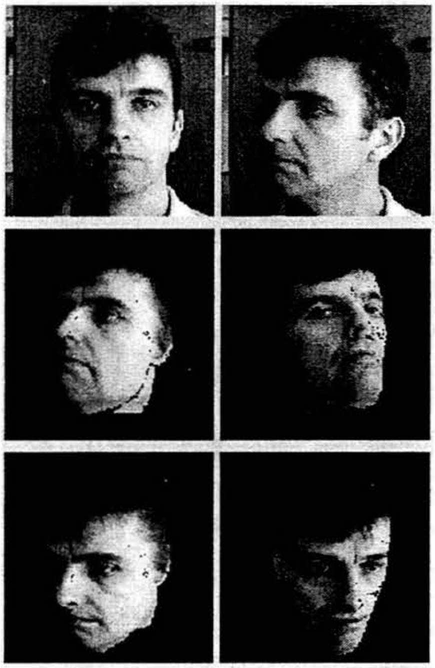
Collaboration:

Future Directions: The current prototype is complete.

Contacts: Paul Beardsley March 13, 2000

Reference: http://www.merl.com/reports

Building 3D Models of the Human Head



See Color Plate #9

The goal of this project is to generate 3D models of the human head. The illustration shows video images of a subject at top. These images are used to automatically construct a 3D model, and views of the model (synthetically generated views of the face) are shown at bottom. This 3D data is useful for a variety of purposes. It can be used for recognition. It can provide the basis for a system that tracks the motion of the subject's head. It can be used to identify expression change. It can be used to generate photo-realistic avatars. Combinations of these technologies underpin significant applications. For example, transmitting a 3D model of a subject's head, and subsequently transmitting just the head motion and expression change, offers significant compression of the data compared to transmitting a full video stream.

Background and objectives: 3D modeling of the human head underpins multiple applications. A prime objective is to investigate the use of 3D models for face recognition from oblique views of a subject (this contrasts with face recognition from fronto-parallel views of a subject, which is now a well advanced area). An example scenario would involve a subject's face being scanned on entrance to an Intelligent Building, and subsequent identification of the subject from elevated cameras at locations inside the building.

Technical discussion: The current system works with a single camera, and utilizes a prior model of a generic human head. The extraction process takes about two minutes and requires controlled head motions by the subject. Newer work is utilizing a three-camera stereo system and has the goal of avoiding the requirement for controlled head motion.

Collaboration:

Future Directions: This technology is of general use for applications that involve observing people, either for surveillance, or for interaction with the subject.

Contact: Paul Beardsley February 7, 2000

Reference: http://www.merl.com/reports

Personal Eyewitness CarCam – Vehicle Accident Video Recorder

The Personal EyeWitness Vehicle Accident Video Recorder provides a robust and tamper-resistant recording of vehicle accidents. The PEW continuously records video into semi-conductor memory, overwriting old video every thirty seconds, and stopping only after an accident triggers the vehicle crash sensor.

The PEW uses Mitsubishi's Artificial Retina chip as an inexpensive image sensor, and Mitsubishi's M32R/D integrated CPU plus DRAM chip to provide data compression and storage. The entire device will be able to sell for approximately US$100 retail.

PEW – Personal Eye Witness, AR: Mitsubishi's CMOS-based Artificial Retina imaging sensor.

Background and objectives: The Personal EyeWitness project goal is to produce a working prototype of a self-contained solid-state video recorder for vehicle accident data capture. The justification for this work is the large market in the US for any system that decreases liability or insurance costs for automobiles, taxicabs, trucks, and buses.

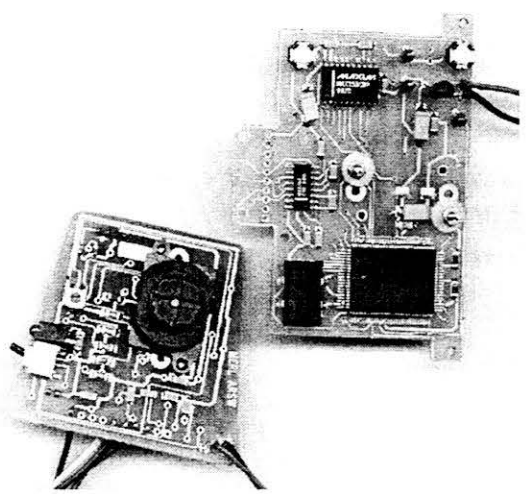
Technical discussion: We have constructed twelve prototype units of the CarCam using the M64283FP Artificial Retina image sensor, the M32R/D microprocessor with 2 MB of DRAM, a solid state accelerometer and 4 MB of FLASH memory. The units are each about the size of a pack of cards and are impact resistant. For demonstration purposes, the units can be attached to small radio-controlled trucks, which are then used to simulate crashes.

Customer focus: The prototype units have been demonstrated at two large shows that Mitsubishi Electric Automotive America put on for Daimler-Chrysler and General Motors. The hands-on nature of the demonstrations as well as the "fun" aspects of camera-equipped toy cars proved to be a very popular combination and the CarCam demo was regarded by many as the "hit" of the shows. The units will be demonstrated to Ford at another large show this September. We are also investigating security applications of the CarCam technology and have some interest from the manufacturer and distributor MCM.

Future Directions: Since most of the technical issues have been worked out, our goal is now to find a customer willing to license our IP and technology. In the mean time, because of the amount of processing power in them, the prototype units make an excellent test bed for portable "smart" camera applications.

Contacts: William Yerazunis, Darren Leigh　　　June 22, 2000

PalmCam – Digital Camera for PDA



The PalmCam is an experiment in using Mitsubishi Electric artificial retina (AR) chips as an accessory for a Personal Digital Assistant (PDA).

The PalmCam experiment explores the use of a highly portable digital camera integrated into a personal digital assistant.

The PalmCam uses an M64283FP AR chip as an imaging sensor, with an M16C microcontroller to produce the control signals, perform A/D conversion, and interface with the PDA serial communications line.

We have developed PalmCam units in several form factors, to fit different PDA applications.

PDA: Personal Digital Assistant, AR: Artificial Retina

Background and objectives: Increasing use of digital imaging integrated into communications and messaging will provide an increasing market for Mitsubishi imaging chips.

Technical discussion: The PalmCam is a generalized design for an imaging accessory for PDAs and other consumer electronics. PalmCams can be implemented in different form factors (two are shown in the photo above). PalmCams use a Mitsubishi Electric AR chip to sense an image, and a Mitsubishi M16C microcontroller to generate control signals for the AR, to convert the sensed image from analog voltage to a digital form, and to interface to the PDA or other consumer electronic device. To decrease costs, the PalmCam uses the on-board A/D converter built into the M16C. Additionally, the M16C can perform image processing on the image or produce a "movie".

Collaboration: PalmCam was developed in cooperation with ARPro in Japan, and with cooperation of the DECWET applications group in the USA.

Future Directions: The PalmCam has been protected by four US patents, and we are actively exploring the profitable licensing of this intellectual property to PDA companies.

Contacts: William Yerazunis, Darren Leigh June 23, 2000

SCAR – Super Cheap Artificial Retina Evaluation Board

The design goal of the SCAR board is to produce and market an AR evaluation board that can be sold at a single-unit price of under $100 US. This goal is achieved by requiring only the absolute minimum hardware, by using the parallel port of a PC as the controlling interface, and by generating all control signals in software.

The SCAR software runs under both LINUX and Windows 98, is open source, and allows all operating modes of the M64283FP AR chip to be evaluated by a system designer, including the powerful image processing modes that are our strongest marketing differentiator.

SCAR : Super Cheap Artificial Retina AR: Artificial Retina (Mitsubishi's CMOS imaging sensor)

Background and objectives: The goal of the SCAR board project is to develop a hardware board and controlling software that can be used by MELCO's USA marketing to leverage the MELCO AR chips into US markets by generating design wins with a low-cost evaluation board. We have beta-tested fifty SCAR boards in the US market, with a strongly positive response, and have been negotiating the sale of SCAR boards through a quick-response distributor in the USA. We hope the SCAR board produces strong growth in the US market for Mitsubishi AR chips.

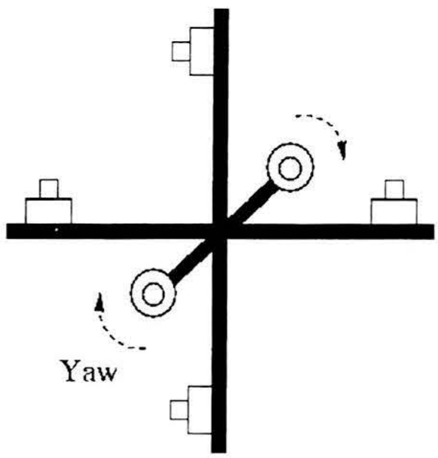
Technical discussion: In order to absolutely minimize the hardware cost of the SCAR board, only the minimum components are included, such as the AR itself, an inexpensive plastic lens, a pair of buffers, and an analog-to-digital converter. Power is supplied externally from a 9-volt battery. To minimize costs, no case is provided, and the circuit card is laid out without multi-layer techniques. The current design uses the M64283FP AR chip. All control signals for the AR are generated by running a program on a PC, under either LINUX or Windows 98. The PC program generates the timing waveforms in software, triggers the A/D converter, and captures the digitized image. The result is that software (which has zero per-unit cost and can be distributed by the Internet) replaces hardware devices that would increase per-unit price. In beta tests, we have achieved speeds of up to 7.5 frames/second with modern PCs. All of the image-processing modes of the AR chip are accessible, including edge detection, image enhancement, convolution, and baseline projection.

Collaboration: This work was done with the assistance of the ARPro group.

Future Directions: We are currently negotiating final distribution and sales details for SCAR boards.

Contacts: William Yerazunis June 1, 2000

Motion-Based Optical Sensing with Multiple AR Cameras



In this project, we investigate novel designs and algorithms for an optical sensing device which can be used as a 6-DOF "joystick" and a servo/controller for real and/or virtual objects. The optical motion sensor makes use of MELCO's own "Artificial Retina" (AR) CMOS chip, an inexpensive, fast and low-power image sensor whose on-chip processing can speed-up the computation of 2D camera motion. By attaching these motion-sensing units to an object (human, toy, car, etc.) we can measure its (relative) motion in real-time using passive optics. Applications include: an optical mouse (2D pointing device), an optical "wand" (3D pointing device) and servo control of real/virtual objects.

Abbreviations: DOF: Degree(s)-of-Freedom, AR: Artificial Retina

Background and objectives: In computer vision, the recovery of camera motion (and 3D scene structure) from optical flow is a relatively mature field and is the basis for various passive navigation techniques. Assuming that the scene is motion-less, the optical flow will be due entirely to the 3D motion field resulting from the camera movement (egomotion). However, we propose a much simpler analysis: given minimal occlusion, large depth, little specularity and minor lighting changes, the optical flow field of ego-motion can be represented globally by a single displacement vector which summarizes the camera motion as projected on the image plane, referred to simply as the 2D motion. Multiple 2D motion sensors of this simple design, when properly attached to a moving object, should recover all 6 DOF necessary for tracking.

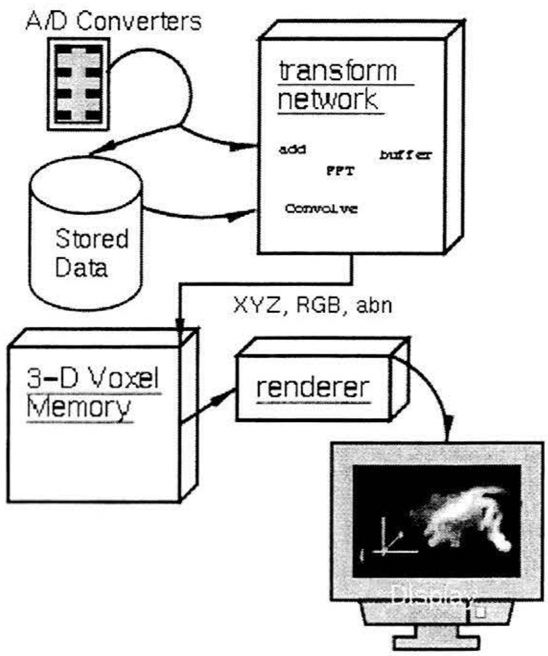
Technical discussion: One simple but effective strategy is to compute the displacements independently using horizontal and vertical projections of the image. The resulting 1D projection signals can be easily tracked with inter-frame analysis (displacement matching) to compute the 2D motion vector (dx,dy). MELCO's own "Artificial Retina" CMOS image sensor ("AR chip") is well-suited to this task since it can compute horizontal and vertical projections in hardware ("on-chip"). The most elementary 6-DOF sensor consists of 3 orthogonally mounted cameras with a camera looking outward along each axis. In theory, one can recover all 6 DOFs with this configuration. In practice, redundancy (more than 3 cameras) enhances performance.

Collaboration: Bill Freeman, Eberharde Lange (Sanken)

Future Directions: DOF recovery using Kalman filters would be a standard approach, but due to ambiguities (e.g., lack of depth) we propose a re-formulation with Bayesian Belief Nets which are better at managing uncertainty and constraint propagation between sensors with unreliable data. Ultimately we would like to test our ideas with a hardware prototype "3D mouse".

Contacts: Baback Moghaddam September 1, 1999

Hyperdimensional Oscilloscope



The hyperdimensional oscilloscope allows a user to examine data in an arbitrary set of dimensions, which may be altered and recomputed interactively by the user. Unlike a normal oscilloscope, the hyperscope considers the time "dimension" to be just another element in a data element. The result is that the user can examine the data for correlations and situations previously unknown (such as emergent behaviors, unsuspected modes, and acausal relations).

Background and objectives: The hyperdimensional oscilloscope is designed as a general-use tool in a large number of areas involving data of high dimensional count, such as medicine, distributed computation, Internet monitoring, public utilities, transportation, weather, and financial and economic modeling.

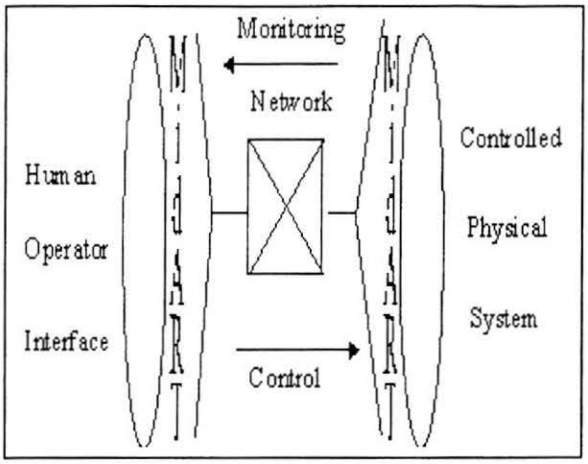
Technical discussion: Inputs can come from both analog and digital sources, and from both real systems and simulations. The data elements are then manipulated by the interactive transform network, which provides an arbitrary user-defined set of computations, on the incoming data set. If the computation indicates that the event being monitored is of interest to the user, the data element is transformed into an RGB tuple, to be plotted at location XYZ, with optional alpha blending and transparency, voxel data fusion, and multiple visual frames, into the 3D voxel memory. The voxel memory is then rendered by either a hardware renderer such as the Mitsubishi VP500 volume graphics chipset, or by a software renderer. Because there is no preconcieved notion of time as a special dimension, the hyperdimensional oscilloscope can manipulate time-based evolutions of systems such as may be needed in intelligent transportation systems, public utilities, or medical, biological and ecological processes.

Collaboration: We are initially collaborating with the RTViz group within MERL, with Cambridge Systems weather and cartographic effort, and with the SanKen weather-radar group.

Future Directions: The hyperdimensional oscilloscope is currently in research design stage; we have detailed designs of about half of the system and are working to complete a testable prototype for human factors analysis.

Contacts: William Yerazunis, Mike Walterman June, 2000

MidART: Middleware for Distributed Real-Time Systems



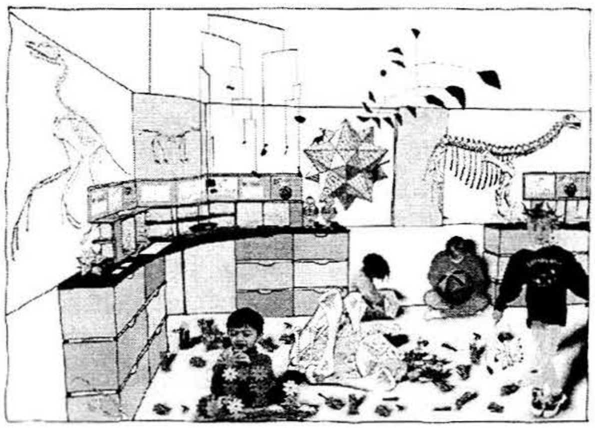
MidART is a distributed real-time software package with easy-to-use programming interface for multimedia, data acquisition and communication. It allows rapid development of multi-threaded concurrent communicating real-time applications in a network environment. MidART Version 1.0(NT) for Windows NT 4.0 and Version 1.0(POSIX) for Unix platforms have both been released. So far, eight universities have signed MidART source license agreement to use MidART in various research projects. MidART has also been included in Sanden's future industrial plant control products.

Background and objectives: Existing network software facilities such as the socket interface are difficult to use for application builders. Moreover, real-time applications need end-to-end quality of service provision. To facilitate the construction of distributed real-time applications on open off-the-shelf systems, there is a strong need to first provide easy-to-use real-time programming models and services to real-time application designers. MidART fills this need.

Technical discussion: MidART provides a set of real-time application specific but network transparent programming abstractions that support individual application's data acquisition, communication and QoS requirements. The focus of the middleware is to support the end-to-end application real-time data transfer requirements with a set of easy-to-use communication service programming interfaces. The key services provided by MidART are Real-Time Channel-based Reflective Memory (RT-CRM) and Selective Channels. RT-CRM is a software-based reflective memory similar to the producer/consumer model -- it provides data reflection with guaranteed timeliness. Data reflection is the memory-to-memory data transfer between application host memory in a networked environment. Many communication models are supported by RT-CRM including the traditional one-to-many and one-to-one models, as well as a unique many-to-one model which is especially useful for many real-time monitoring and control applications where one or few operator stations would need to keep track of and control many (e.g., tens or hundreds) devices and instruments. With MidART, the application designers can specify how and when the data should be sent according to the application's specific needs. RT-CRM achieves data reflection with an underlying active data push mechanism. Various data push and data retrieval modes are supported, including synchronous vs. asynchronous data push, and blocking vs. non-blocking data retrieval. These modes can be effectively combined to achieve many kinds of application specific tasks. Selective Channels allow applications to dynamically choose the remote node(s) which data is to be viewed from and sent to at run time via a set of channel start and stop protocols, and channel bandwidth resource reservation schemes.

Collaboration: MELCO's IESL and University of Massachusetts at Amherst.
Future Directions: We are working on making MidART a standard CORBA real-time service.
Contact: Chia Shen
Reference: http://www.merl.com/projects/midart June 27, 2000

Scientific Studio: A Model for Museum-based Interactive Learning Environments



See Color Plate #10

Abbreviation: SciStudio

The model of a studio combines a lab-like function supporting experimental production, a gallery function supporting participants' display and discussion of what they produce, and an apprenticeship structure in which learners work side-by-side with skilled practitioners. The environment bridges the physical and virtual worlds by providing design environments and dynamic simulations in the virtual domain, supporting construction of the designed objects as crafts in the physical domain, and enabling communication between the two domains.

Background and objectives: We are addressing needs for new ways of presenting math and science, and new kinds of environments in which learning about math and science can happen, through development of a model for environments that rely on computer technologies and a learn-by-doing philosophy to support participants of many ages and thinking styles.

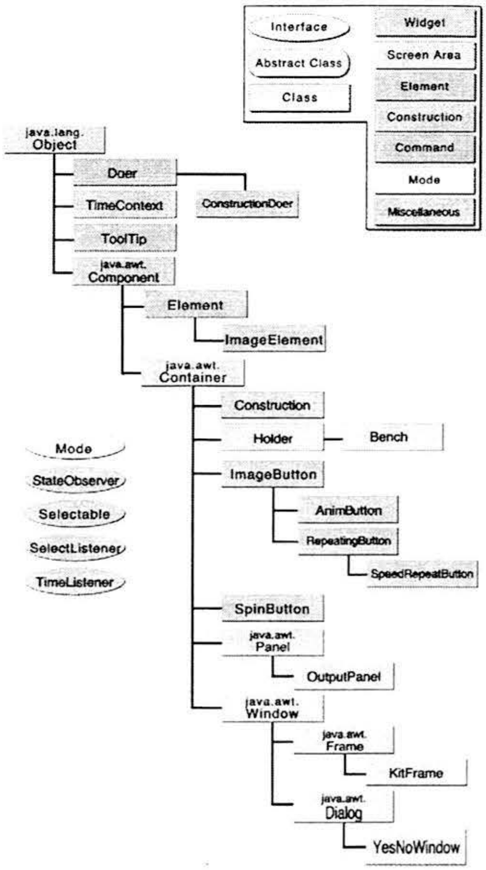
Technical discussion: The learning environment includes both physical and virtual spaces, each of which functions as a work area and an exhibit area. Outside suppliers provide construction materials and learning tools, many of which are available through online museum stores. Activities in the physical domain include design, construction, and display of craft items. Craft construction may be assisted by materials with embedded computational capabilities, and by software kits with tangible interfaces and unusual devices facilitating inputs and outputs. The physical environment supports tinkering individually or building collectively while surrounded by constructions at various stages of completion. The physical space also includes large-screen displays that provide views into the virtual domain. The virtual world includes software construction kits and a multiparticipant domain for show and tell about dynamic creations. Again, activities include design, construction, and display of objects, but with a focus on aspects that would be difficult or impossible physically. These aspects include changes of scale, multiple simultaneous views, animation, simulation, and remote presence.

Collaboration: The collaboration includes MERL Cambridge Research, the University of Colorado at Boulder, Boston's Museum of Science, and some people at Horizon Systems and MELCO's Shiden lab.

Future directions: In time the environment should be distributed across several museums and many participants' homes, so that experiences can continue for longer periods of time than a single museum visit would allow. Meanwhile we need to develop examples of computational crafts, tangible I/O for software kits, and communications infrastructure to support participants' display, discussion, and exchanges of constructions.

Contact: Carol Strohecker February 25, 2000

The Kit4Kits: A Java Framework for Software Construction Kits



The Kit4Kits is a Java framework for software kits that are highly visual and highly interactive, and encourage learning-by-doing. The implementations of four existing kits (Bones, WayMaker, PatternMagix, and AnimMagix) have informed development of the framework, which we are refining as we apply it to development of a new kit (Zyklodeon). The Kit4Kits includes code for specifying, tracking, altering, and reporting on system states; for creating structure, function, and appearances of objects; for generating screen layout items and widgets; and for capturing and dealing with actions in event-driven systems. Users of the Kit4Kits are learning researchers, interaction designers, and Java hackers who create kits for end-users – who in turn use the kits to build graphical objects that come to life on the screen. The framework includes guidelines for conceptual grounding of content and strategies for kit formation and interaction design.

Abbreviation: K4K

Background and objectives: The Kit4Kits supports development of kits within a genre premised on constructivist approaches to learning and learning research. End-users create two-dimensional, graphical forms that behave in characteristic ways when activated. Dinosaur skeletons balance as they walk and run; maps reveal street-level views; geometric tiles form symmetric patterns; animistic creatures establish social distances; dancers' breathing determines timing for a shared dance; and so on. Our kits address conceptual domains of topology, geometry, symmetry, sensori-motor functions, time/space relationships, and system dynamics.

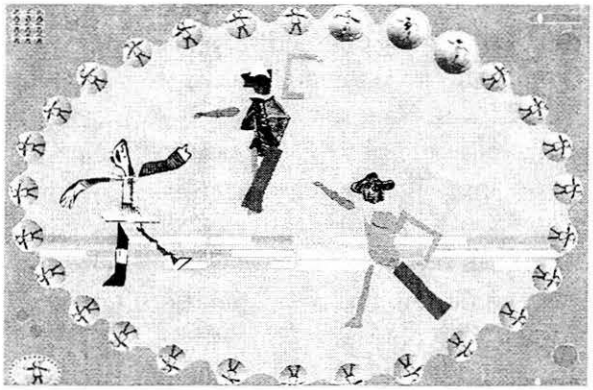
Technical discussion: The Kit4Kits includes structures that extend the Java Abstract Windowing Toolkit (AWT). Among them are highly customizable screen areas, widgets, and image treatments such as painting techniques, geometric transformations, transparencies, and color blending. Code examples and "seeds" maximize accessibility for beginning implementers. Design heuristics promote multiple simultaneous views; miniature representations of temporarily inactive screen areas; visualizations of algorithms and processes; and visual and aural feedback.

Collaboration: Carol Strohecker and Adrienne H. Slaughter designed and developed the Kit4Kits. John Evans, Aradhana Goel, Tim Gorton, and Milena Vegnaduzzo participated in a preliminary usage session that helped to focus and refine the framework.

Future directions: The Kit4Kits is part of a larger effort to develop museum-based interactive learning environments. The effort involves MERL Cambridge Research, MERL Cambridge Systems, MELCO's Shiden lab, Boston's Museum of Science, and the University of Colorado at Boulder.

Contact: Carol Strohecker February 25, 2000

Zyklodeon: A Kit for Thinking about Time and Dynamics



Zyklodeon is a software construction kit that supports creation of colorful figures whose combined breathing rates form a cycle for a shared dance. The dancers' movements are visualizations of progress through the cycle and of characteristics that affect choreographic concerns such as leap moments and heights. In this artistic world, learners experiment with scientific principles such as time/space relationships and emergent properties of dynamic systems.

Background and objectives: Zyklodeon is part of a genre of software construction kits based on the idea of "microworlds." These carefully designed learning environments emphasize construction of objects as a parallel to construction of ideas and understandings. Microworld users manipulate representations to focus thinking on particular aspects of a conceptual domain. The primary domain of Zyklodeon is motion study. Dancers in the Zyklodeon move with respect to one another, according to cyclic timing parameters that derive from the figures' own internal rhythms. The secondary conceptual domain of Zyklodeon is system dynamics, which is a method for studying how complex systems change over time. Zyklodeon users can experiment with changes of breathing rate and other parameters to see how one variable can influence the behavior of an entire system.

Technical discussion: Zyklodeon is implemented with MERL Cambridge Research's Kit4Kits Java framework. We include main screen areas for composing figures, watching them dance, and storing them. The user assembles a dancer by choosing body parts to form a six-part figure. Heads, arms, and legs repeatedly separate from the torso and return to it as the figure "breathes." User-changeable settings determine the separation distance, frequency, and movement characteristic, as well as the degree to which the torso can lean as the figure translocates. These settings form the figure's "drama factor." The frequency of a dancer's breathing combines with that of another figure to determine the number of breaths per dance cycle. This cycle is the timing regulator for the figures' shared dance. As it plays out, the drama factors increase. When a figure's drama factor reaches a certain threshold, the figure leaps. The drama factor then resets and the figure resumes its movements during the dance cycle.

Collaboration: Carol Strohecker, Adrienne H. Slaughter, Michael A. Horvath, and Noah Appleton designed and developed Zyklodeon.

Future directions: Zyklodeon is part of a larger effort to develop museum-based interactive learning environments. The effort involves MERL, Horizon Systems, MELCO's Shiden lab, Boston's Museum of Science, and the University of Colorado at Boulder.

Contact: Carol Strohecker February 25, 2000

MERL Cambridge Systems

MERL Cambridge Systems is poised to be a prime contributor in an age of accelerating digital convergence, ubiquitous computing, and increasing connectivity and user mobility. To meet these challenges, the Lab is developing innovative technologies: systems software (such as operating system or kernel software), middleware for communications or application specializations, and unique applications as well as productivity tools. This range of development items is provided by an experienced and highly skilled staff with over 300 cumulative years of experience in product development and a leading edge laboratory infrastructure with an excellent support staff.

MERL Cambridge Systems builds solutions either directly or by developing key enabling technologies and tools. We focus on three major solution domains:

- **Intelligent Virtual Enterprise**: the full range of wired and unwired connectivity and integration, from the sales force and E-commerce to the factory floor. MERL Cambridge Systems provides many advanced technologies and software tools in this solution domain: Concordia, Concordia EI, Scaleable Mobile Computing, and NetRep. This effort also focuses on expanding the modularity (plug and play capabilities), extensibility, and scalability of next-generation enterprise systems in order to accommodate the potpourri of new computing paradigms and associated hardware devices.
- **Social Computing**: the real-time connection between people for social applications, such as collaborative working environments, gaming, and networked social spaces. Here, the Lab provided Open Community, and Schmoozer.
- **Interactive Architecture**: a platform that enables security and services in commercial/residential buildings and public spaces. This effort combines MERL Cambridge Research's expertise and IP in vision and real time networking, MERL Murray Hill's advances in communications protocols and standards, and MERL Cambridge Systems' experience in system engineering, human computer interfaces, and ubiquitous environments. Related technologies are computer vision applications, MPEG standards, and Concordia.

MERL Cambridge Systems works in cooperation with MELCO business units, affiliates, and laboratories on a number of efforts. Within MELCO, the lab cooperates closely with the Social Infrastructure and Information business units, the Industrial Electronics & Systems Laboratory and the Information Technology R&D Center. In addition, the lab partners with U.S. organizations for both technology and business opportunities, including established software companies such as Veritas Software and small technology innovators such as ActiveIQ.

Due to its product-oriented nature, MERL Cambridge Systems' output has had a direct impact on MELCO's bottom-line. The lab has the distinction of having received two MELCO President's awards and two Corporate R&D General Manager's Awards, both unprecedented for an overseas advanced development organization.

To allow better collaboration between MERL Cambridge Systems and MERL Cambridge Research, MERL Cambridge Systems is in the process of moving from a building in Waltham, MA, to MERL's Cambridge, MA, facility. As of the summer of 2000, approximately half of the people have already moved. The move will be complete by the summer of 2001.

Technical Staff

Andrea DeDuck

Principal Technical Staff

Andrea DeDuck has worked on numerous projects in her years at MERL Cambridge Systems as a CPU micro-code developer and systems verification team leader. Most recently she was project leader on the Netrep project, which produced MERL Cambridge Systems' first commercially licensed technology. In her spare time, Andrea likes to visit exotic places above and below sea level.

Susan Deisenroth

Senior Quality Assurance Engineer

Susan has been working as the Senior Quality Assurance Engineer for Concordia Development team. Prior to CSL, Susan worked for Platinum Technology managing a two-person team of a DataWarehouse product called DecisionBase. She has extensive knowledge of relational and non-relational databases as well as a programming background of C. Prior to Platinum Technology, Susan worked as a software developer for AICorp, a Knowledge Base Management company. There she designed and implemented a number of database programs, written in C, which retrieved and stored data to various databases.

Alan Esenther

Principal Technical Staff

Alan Esenther enjoys Internet technologies and usability embellishments. His most recent work involves automatic web-page generation and an email gateway related to disconnected web access for an Enterprise Application Integration project. He started out doing microprocessor development at Data General and continued with board-level development at MERL Cambridge Systems. Since then he's turned to software development, applying an M.S.C.E. and technical certificates to transaction monitor and kernel-level volume management projects. Interests include rock climbing, adventure travel, and learning new technologies.

Frederick J. Igo, Jr.

Senior Principal Technical Staff

Fred Igo's professional interests are in software development and its process. Starting at IPL Systems and continuing at MERL Cambridge Systems. Fred developed and managed teams developing wide-word microcode and tools as part of mainframe development projects. Two of his four mainframe projects won MELCO's President's Awards. Fred next moved into database and middleware development. During three projects he has worked with DCE, Distributed OLTP message queuing, MELCO's hardware database accelerator DIAPRISM (SPG) and OLAP databases like Oracle Express. Recently Fred has transitioned to Java development, developing applications and tools for MERL Cambridge Systems' mobile agent system Concordia.

Masataka Kawaguchi

Senior Principal Technical Staff

Masa Kawaguchi joined MERL Cambridge Systems three years ago from MELCO. His current interests are mobile agent technology and its applications, developing Java mobile agent middleware for the Concordia project. He has worked on a number of software projects as a project leader at MELCO and has a wide range of software background, from real-time software to application design, as well as project management experience. He also won MELCO awards several times. This is his second stay at MERL Cambridge Systems.

Brenden Maher

Senior Technical Staff

Mr. Maher received his Masters in Media Arts and Sciences from the MIT Media Lab in 1998. His interest in computers began through developing Immersive Real-Time 3D Virtual Environments with the Boston Virtual Reality Group in 1992. With the group, he participated in the design and development of Virtual Environments for SIGGRAPH '93,'94,'95 and '97. As well, he was Director of the group since '94. At the MIT Media Lab, his research involved developing a system for navigating an audio database quickly and efficiently using 3D Spatialized Audio and graphics. As part of the Lab's Speach Interface Group,he continued his development in real-time virtual environments and distributed systems and also developed new forms of human computer interaction. His Audio Database Navigation System used head and body tracking for information navigation. Mr. Maher studied Tangible Interface design under Hiroshi Ishi and developed a new type of "Ambient Display" called the Bit Blower to provide information in the periphery of one's focus. He has worked in interactive media (CD-ROM's) and built a 3D data visualization system for analysis of Mechanical Engineering Data while at Molecular Geodesics, Inc. Currently, he is involved in the research and development aspects of Mitsubishi Electric- Cambridge Systems Laboratory's ubiquitous computing efforts called Interactive Surroundings.

Jacob Nikom

Senior Technical Staff

Jacob Nikom's research interests center around duality of geometry representation in vector he used in modeling robot kinematics in space with obstacles, working in the Institute for the Machine Studies and at Applicon/Schlumberger Company, and raster forms at Hewlett-Packard Corporation he developed computer vision applications for MRI and ultrasound imaging modalities and at the University of Massachusetts Medical Center doing 3D image-processing research. At Mitsubishi Electric he has worked on real-time physical simulators for angiography procedures and currently implements computer vision algorithms for Artificial Retina chips.

Noemi Paciorek

Senior Principal Technical Staff

Noemi Paciorek has worked on several projects over her six years at MERL Cambridge Systems and is currently a key member of the Concordia core development team. Prior to joining MERL Cambridge Systems, Noemi participated in leading-edge technology development in real-time systems, distributed and multiprocessor systems, and operating systems. She has co-authored dozens of papers and technical reports on a broad range of topics and is a recipient of multiple best-paper awards.

Luosheng Peng

Principal Technical Staff

Before joining MERL Cambridge Systems, Lousheng Peng was a senior engineer at Mitsubishi Space Software working for MELCO Advanced Technology Lab's Nuro 4 project, and was a active director at Fonte System, Inc. developing a mechanical CAD/CAM system. Luosheng developed a Japanese-government-sponsored intelligent CAD system for that he received three Excellent Paper awards. Luosheng's current research includes data replication and synchronization for mobile computing and application proposals and development.

David C. Rudolph

Principal Technical Staff

David Rudolph has been at CSL for 10 years. During this time he has contributed to several systems software projects, including the gm80 simulator. His last 4 ½ years have been spent developing the "Network Replication" project, which is now being successfully marketed by Veritas Software Corporation as SRVM (Storage Replicator for Volume Manager). This project was the 1998 winner of the Corporate R&D GM's Award for Excellence. Before joining CSL, David spent three years at Data General, interrupted by a two-year stint at the University of Illinois, where his research interests were system software and performance analysis for massively parallel architectures.

Kathy Ryall

Principal Technical Staff

Kathy Ryall received the BA degree with a double major in Mathematics/Computer Science and Psychology from Wesleyan University, and the MS and Ph.D. degrees in Computer Science from Harvard University. Prior to joining MERL Cambridge Systems in July 2000, Kathy served as an Assistant Professor of Computer Science at the University of Virginia for three years. Kathy's primary technical interest is the design and implementation of systems in which the interface acts as a medium for people and computers to work together on solving problems, rather than as a means for people to control computers. This technical interest has resulted in the development of systems for a variety of tasks in the computer graphics and graphic design domains.

Derek L. Schwenke

Principal Technical Staff

Derek Schwenke received his B.S.E.E. from Tulane and M.S.C.S. from Worcester Polytechnic Institute. His work areas have included the computer software and hardware fields. He has worked with image processing and satellite communications systems at Raytheon Corp. in Marlboro, MA. Derek began his work at MERL Cambridge Systems in 1988 with the design and simulation of M80 and PXB1 CPUs, and progressed to software development with the OSF-DCE/Encina system. He co-developed the OMQ message queuing system for MELCO. In the last three years Derek has worked extensively with Open Community virtual reality system developed at MERL Cambridge Research Laboratory in Cambridge, MA. Most recently he's helped develop the Internet Sharing and Transfer Protocol (ISTP).

Samuel E. Shipman

Principal Technical Staff

Sam Shipman's technical interests include real-time and distributed operating systems principles and techniques and their application to distributed virtual reality systems. Sam currently leads the Open Community Core project at MERL Cambridge Systems. Previously, he developed Unix kernel code for the Network Mirror implementation. Prior to joining MERL Cambridge Systems, Sam worked on the Mach kernel at the Open Software Foundation and at the Center for High Performance Computing at Worcester Polytechnic Institute. Earlier, he was a principal developer of Alpha, a real-time distributed operating-system kernel embodying innovative research in scheduling and process management. Funded by the United States Air Force, Alpha was developed at Concurrent Computer Corporation and originated at Carnegie Mellon University, where Sam received a M.Sc. in Computer Science.

Mike Walterman

Principal Technical Staff

Mike Walterman's major professional interest is in the application of real-time visual image generation to medical and vehicular simulation. His background is in image processing, computer graphics, and systems engineering. Before joining MERL Cambridge Systems, Mike worked for five years as a software and systems engineer at Evans & Sutherland. Mike has 18 years' experience working in companies ranging from small startups to multinational corporations. In addition to his normal system development duties, Mike has published papers on image warping techniques, lectured on immersive simulation, and has patents pending on image rendering technology. In his spare time, Mike enjoys mountain climbing, cross-country running, and watching lacrosse.

David Wong

Deputy Director

David Wong's major professional interest is in incorporating mobile agents into mainstream mission-critical applications, spanning traditional business IT domains to industrial automation. His background is in mobile agent systems, transactional message queuing systems, and distributed transaction processing. He has been with MERL Cambridge Systems for the past six years. Prior to joining Mitsubishi, David worked on the advanced development and performance analysis of transaction processing systems at Compaq (the former Digital Equipment Corporation). He has also taught and conducted research at the University of Connecticut and Brown University. In a prior life, David also dabbled in research on theoretical thermodynamics. He holds a Ph.D. in computer science from the University of Connecticut and a B.Sc. in chemical engineering from Brown University.

Michael Young

Senior Principal Technical Staff

Since joining MERL Cambridge Systems in 1996, Michael has been involved in the design and development of the Concordia Mobile Agent system. Michael began his career at Raytheon Company, where he specialized in distributed, fault-tolerant systems and real-time embedded operating systems. He left Raytheon in 1987 to become a founding engineer of a new software development company that sought to capitalize on the burgeoning interest in the C++ language. Immediately prior to joining MERL Cambridge Systems, Michael was an engineering supervisor and group leader specializing in distributed systems and asset management software.

Recent Major Publications

Wong, David., "Concordia Java-based Mobile Agent Systems Framework", *Mitsubishi Electric Technical Journal*, March 2000.

Nikom, Jacob., Walterman, Mike., "ICTS: An Interventional Cardiology Training System", *Proceedings of Medicine Meets Virtual Reality 2000*, pp. 59-65, January 2000.

Wong, David., Castillo, Alberto., "Applying Java-based Mobile Agents to Industrial Automation", *Proceedings of the First Java Technology for Industrial and Societal Automation Conference*, June 1999.

Koblick, Ruvin., "Going Mobile: Mobile Agents Drive Home the Next Generation of Middleware?" *Enterprise Systems Journal*, May 1999.

Wong, David., Paciorek, Naomi., "Java-based Mobile Agents", *Comm. ACM*, Vol. 42, No. 3, pp. 92-102, March 1999.

Koblick, Ruvin., "Concordia", *Comm. ACM*, Vol. 42, No. 3, pp. 96-97, March 1999.

Walsh, Thomas., Paciorek, Naomi., Wong, David., "Security and Reliability in Concordia", *in Mobility: Processes, Computers, and Agents*, pp. 525-534, New York, 1999.

Paciorek, Naomi., Wong, David., "Concordia: A Mobile Agent Platform for Complex Distributed Systems", in *Proc. Agent Technologies Workshop '98*, pp. 24-37, November 17-18, 1998.

DiCelie, Joseph., "Concordia and Its Security Features", in *Proc. Japan Society of Software Technology Workshop*, pp. 133-140, August 13-15, 1998.

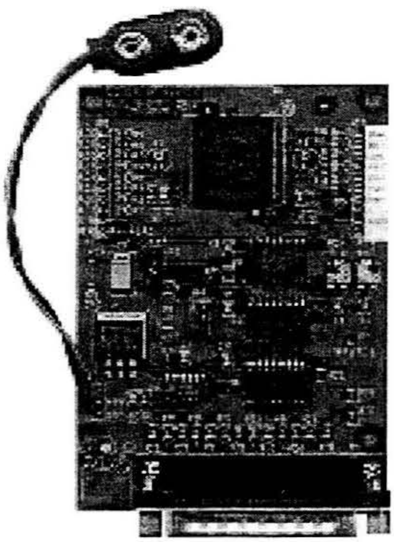
Castillo, Alberto., Kawaguchi, M., Paciorek, Naomi.,Wong, David., "Concordia as Enabling Technology for Cooperative Information Gathering", in *Proc. Japan Society of Artificial Intelligence Conference*, pp. 155-159, July 21-22, 1998.

Peng, Luosheng., " A Star-structured Data Consistency Model for Wide-area Mobile Network Computing", in *The Sixth IEEE Singapore International Conference on Networks*, July 1998.

Walsh, Thomas., Paciorek, Naomi., Wong, David., "Security and Reliability in Concordia", *in Proc. Thirty-First Annual Hawaii International Conference on System Sciences*, pp. 44-53, January 9-11, 1998.

Young, Michael., Wong, David., "Mobile Agents: The Next Generation in Distributed Information Management", *Storage Management Solutions*, Vol. 3, No. 6, pp. 16-20, 1998.

Project Descriptions



Artificial Retina Skunkworks

An "Artificial Retina Skunkworks" project started on April 2nd, 1999. The reason for the project was it would increase the sales of the AR chip (and other Melco chips such as the M32 and M16 series) by developing prototype systems utilizing these chips. **Prototypes developed:** 1. Elevator People Counter – counting people at elevator entrance. 2. Proximity Detector – measures distance to the moving object based on how far the object is out of focus of the camera. 3. Camera Motion Gauge – translation and angular camera motion measurement using built-in features of AR chip to speed up the calculations.4. LegCam – measuring the distance between vending machine and people based on monitoring people's legs activity.
5. Image Acquisition Toolkits: Matrox, WebCamII, Artificial Retina and Pamette Stereo Board.
6. Java-based graphical user interface for stereo video image acquisition and analysis for automatic detection of blind spot of a car.

Abbreviations: AR – Artificial Retina, GUI – Graphical User Interface

Background and objectives: The goal is to locate specific opportunities in the North American market that can be satisfied by vision-enabled ubiquitous computing devices; and creating prototype systems.

Technical discussion: The developed prototypes utilize image-processing-on-chip feature of AR sensor. They perform the very simple task: evaluate simple human activity – walking, wheeling, staying, gathering . People Counter prototype could be used for counting people in lines, LegCam – to count customers who did not buy the product but was close enough to be considered as "looking to buy". Camera Motion Gauge could be used for such application like optical mouse. Java-based GUI for the stereo imaging algorithm development for the car side blind spot detection system allowed viewing and manipulation of up to 12 images(multiple stereo), flexible positioning of images inside the GUI, ability to work with images sequences as well as single images (acquisition, viewing, analysis, storage), ability to draw geometry (points, lines and polygons) over still and video images in real-time, real-time report about pixel properties under the cursor.

Collaboration: MERL Cambridge Research groups: three principle hardware prototypers; four computer vision researchers. MELCO: ARPro organization in Japan, which is responsible for the AR Chip line.

Future Directions: Commercialize developed prototypes with ARPro, Sunken and "Aging at Home" project. Add Internet connectivity to image acquisition toolkits.

Contacts: Jacob Nikom 28 April 2000

Concordia

The ***Concordia*** Java-based mobile agent systems framework was developed to address the needs of the mobile user. ***Concordia*** is the most comprehensive product among commercial offerings and can be tailored to fit the particular hardware needs of the mobile user due to its modular means of deployment, i.e., only the required components need to be installed. ***Concordia*** can be deployed on a spectrum of hardware devices, from smartphones and PDAs (Personal Digital Assistants) to high-end back-room servers to fulfill the needs of mobile users for enterprise wide computing.

Background and objectives: The ***Concordia*** Java-based mobile agent systems framework was developed to address most of the needs that earlier systems such as General Magic;s ***Telescript*** could not provide for the mobile user. The goal of this project was to provide complete mobile agent systems support for the mobile user within the context of enterprise-wide computing.

Technical discussion: ***Concordia*** is the most complete mobile agent systems framework available among all commercial products and research prototypes in the Java-based mobile agents space. ***Concordia*** offers complete systems reliability for agent communication, execution, and transmission, and server robustness in the form of seamless restart and recovery upon system failure. Furthermore, ***Concordia*** offers the most complete security among all mobile agent systems offerings. Its security support provides protection of the agent from tampering by other agents, protection of access to server system resources by unauthorized agents, and protection of agents during transmission via encryption techniques. ***Concordia*** is highly scalable and can be deployed across a spectrum of hardware devices, from smartphones and PDAs to high-end backroom servers; its memory requirements can be tailored to a particular device and solution by deploying only those components that are needed by a particular customer. ***Concordia*** currently comes in two distinct flavors: the Full Server version and a Lightweight Server version (targeted for embedded devices and requiring under 3.0MBs of memory for the Concordia Server).

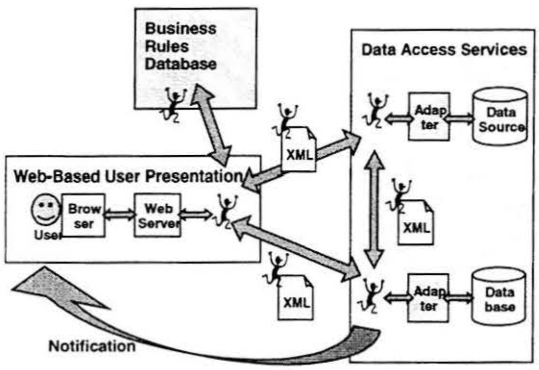
Collaboration: The following organizations within MELCO have used or are currently using ***Concordia***: the Power Systems Business Unit for inclusion in the Osaka Waterworks Maintenance Systems rollout as well as new prototyping projects power companies and the Business Internet Division for inclusion in the ***MELBA*** helpdesk framework. ***Concordia*** has also been used in academic research projects at the following universities: Washington University (St. Louis, Missouri), City College of New York, University of California at Riverside, University of Cyprus, and Cornell University.

Future Directions: Current ***Concordia*** development has already addressed most of the shortcomings of competing mobile agent frameworks. Future R&D efforts would focus on developing a more layered architecture so that ***Concordia*** can be customized for a broader spectrum of hardware devices and underlying software platforms.

Contact: David Wong June 22, 2000

Concordia XML Framework

Concordia XML Architecture



Concordia XML is an environment for the rapid development and deployment of EAI applications. The ***Concordia XML*** system builds upon ***Concordia*** and uses mobile agents to efficiently and flexibly link together enterprise applications and databases. At its core the ***Concordia XML*** system uses XML to encapsulate enterprise information in a standard form. ***Concordia*** mobile agents route XML documents throughout a corporate network based on a customizable set of *business rules* and access back-end data sources via customized system provided *data access adapters*.

EI: Enterprise Integration, **EAI**: Enterprise Application Integration, **SI**: System Integration, **XML**: eXtensible Markup Language.

Background and objectives: The ***Concordia XML*** project grew out of the original ***Concordia*** project with the intention of building technologies that could be used to support an EAI systems integration business. The goals for the project was to build technology that would allow for the rapid development of customized integration applications.

Technical discussion: The ***Concordia XML*** system is made up of *design-time tools* for the rapid development of customized integration applications and of a *run-time environment* for the deployment and execution of these applications. The system builds upon ***Concordia*** and is written entirely in Java. The system operates by converting enterprise data from proprietary back-end formats into industry-standard XML format. ***Concordia*** agents then route this information around a network, manipulating the information and sharing it with whatever systems are required to complete a desired task.

The system's back-end is composed of *data access* to convert the information contained within legacy data source into XML format. On the front-end, the system uses a *web-based user presentation layer* that allows end users to view and manipulate information via a web browser. The system's middle tier is composed of a flexible *business rules architecture.* The system provides design time tools which allow a systems engineer to define how legacy information is mapped into XML format, allows for the automated generation of HTML pages for user presentation, and allows for performance and load simulation.

Collaboration: The following organizations within MELCO have adopted the ***Concordia XML*** tools: power systems business units and IT systems integration division.

Future Directions: Future developments will extend the system by providing generic application templates that can be quickly customized for a specific customer's needs.

Contact: David Wong June 22, 2000

Interactive Surroundings

The **Interactive Surroundings** project builds on earlier Smart HabMERLt and AgingAtHome project efforts. The goal is to build systems which provide better security, safety, and user responsiveness and services in commercial and residential buildings, and in public spaces. This initiative is being developed using core MERL IP in areas of vision, networking and human computer interaction; MERL Cambridge Systems looks towards supporting MELCO Business Units in designing systems and enabling its sales of systems.

Background and objectives: MERL recognizes the need to utilize prior investments in IP and leverage its core competencies. At this stage, the project leverages the prior investments in IP at MERL/ATL. It also leverages MERL Cambridge Systems core competencies in systems integration, software engineering and HCI (Human Computer Interaction). We plan initially to demonstrate technologies that are central to a broad variety of applications -- whether in residential buildings or commercial, or for elderly people or young, or passive surveillance or active interaction. We will then use that technology to support the more specific needs expressed by Business Units.

Technical discussion: Advanced systems will be developed using ubiquitous computing constructs of context of information, and advances in MERL's core competencies of vision, networking, machine learning, human computer interaction and systems engineering. Potential applications include using existing MERL IP in face modeling/recognition, MidArt Real Time Networking, Artificial Retina and machine learning.

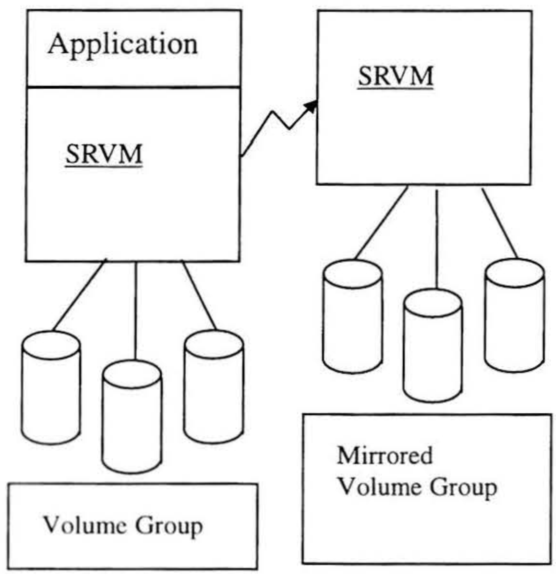
These competencies will be applied toward systems which range in functionality from aiding facility management professionals to end consumer services. Hardware resources will include-cameras and other sensors, multiple computing systems and wireless/networking infrastructures.

Collaboration: The Interactive Surroundings project is a collaborative effort within MERL. The following BU's are potential targets: Buildhon, Shakai INFRA, Tsu Hon,Johon, INethon, Denshihon, Li Hon, Sentan Soken, ID Ken, and Mitsubishi Electric Life Service Corp.

Future Directions:
Future directions of the Interactive Surroundings Project involve advanced development of MERL's IP and Project definition in MELCO's interests.

Contact: Brenden C. Maher June 16, 2000

Network Replication



The Network Replication project is a cooperative effort with VerMERLs Software to provide network replication capability to their popular Volume Manager product, VxVM. Our technology has been licensed by VerMERLs for their Storage Replicator for Volume Manager product (SRVM).

The replication engine of SRVM offers an ideal solution for the remote archiving of storage such as databases or file systems through its ability to replicate an unlimited number of related data volumes while maintaining consistency. Replication uses standard networks without proprietary hardware and is highly resistant to system and network failure.

VerMERLs Software is a market leader in storage management software.

Background and objectives: The Network Replication project evolved from an R&D exploration of network distributed storage. Establishment of our relationship with VerMERLs Software refined this objective to network replication of VerMERLs' logical volumes.

Technical discussion: VerMERLs' customers use VxVM to protect their data from media failure by creating local mirrors, but remote mirrors were needed to protect against system and infrastructure failure. SRVM replicates groups of logical volumes, maintaining consistency among them during replication. Changes to any member of the volume group are transparently captured and replicated to one or more remote locations.

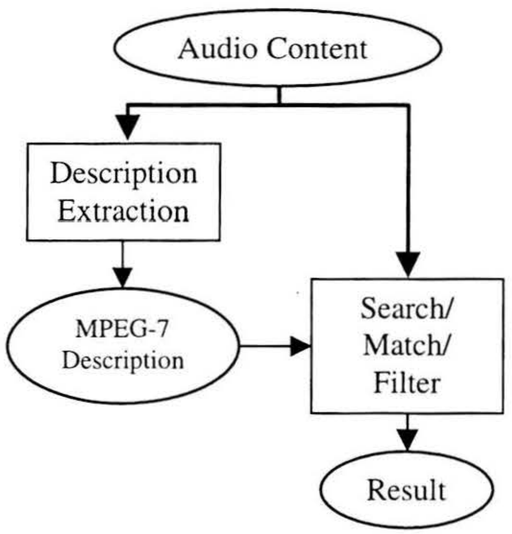
SRVM has flexible replication and configuration characteristics. It can replicate synchronously or asynchronously. Feedback allows input flow rates to be throttled if necessary to match available network bandwidth. It can simultaneously replicate volume groups to remote sites while acting as a receptacle for volume groups replicated from other sites and it can support an unlimited number of volume groups. It can replicate a volume group to multiple destinations, each with independent replication characteristics and latencies

Collaboration: Since 1997, MERL Cambridge Systems has worked exclusively with VerMERLs Software in the development of SRVM, deriving market direction and market access from this relationship.

Future Directions: The base SRVM product will be ported to new platforms such as Windows NT and HP/UX, and its feature set expanded to better serve the disaster recovery market.

Contact: David Rudolph April 28, 2000

MPEG-7 Software Integration of Sound Spotter Technology



Sound Spotter is a technology for analyzing and identifying distinct sound sources in a mixed audio track. MPEG-7 is a standard for annotating multimedia content with machine-readable descriptions.

Sound Spotter generates and uses descriptions that can readily be expressed within the framework of the MPEG-7 standard. Incorporation of Sound Spotter description methods into the MPEG-7 standard would facilMERLte widespread adoption of the technology, enhancing its general utility and revenue potential.

The objective of this work is to perform the system integration and software development required to support MPEG-7 standardization of Sound Spotter core technology.

XM: eXperimental Model. ICA: Independent Components Analysis. W3C: WWW Consortium.

Background and objectives: See the MERL project report "Sound Spotter: Recognition and Extraction from Mixed Audio." Sound Spotter will likely emerge as an important enabling technology for such applications as automated audio content description, searching, indexing, and filtering (among others).

MPEG-7 is a nascent ISO/IEC standard for a "Multimedia Content Description Interface" intended to support as broad a range of applications as possible, including information retrieval, filtering, summarization, etc. To that end, MPEG-7 has specified a Description Definition Language (DDL) based on the W3C's (rapidly evolving) XML Schema specification. MPEG-7 will standardize the DDL and a foundational set of "description schemes" (specified in DDL). Candidate description schemes must be integrated into MPEG-7's reference model software (the "XM"), which is the objective of this work.

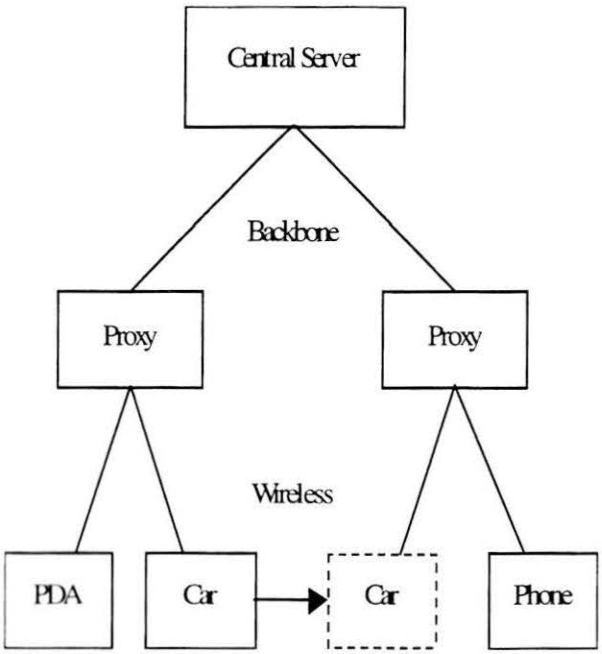
Technical discussion: XM integration requires a precise DDL definition of the ICA Frames description scheme. Software to "extract" the description from audio data, encode and decode the description (using the XML DOM interface), and match similar content descriptions (for searching and indexing) must be integrated into the XM framework. The research prototype of Sound Spotter is written in MATLAB, an interpreted language that can be compiled into library-intensive C++ code, which must be adapted to the XM's pre-existing class hierarchy.

Collaboration: This is a collaborative effort between MERL Cambridge Systems and MERL.

Future Directions: Complete the integration and provide support as needed for standardization.

Contacts: Sam Shipman, Michael A. Casey June 16, 2000

Scalable Mobile Computing



Scalable Mobile Computing addresses the requirements and solutions for mobile computing, a growth area for the 21st century.

Wireless connections will continue for some time to be costly, unreliable, and slow. Supporting cost-effective and continuous operation will require that data accessed by the users must be replicated; in addition, updates that are accumulated on the mobile devices and backbone must be exchanged and synchronized so that they will be kept consistent on all devices.

The project has aimed to discover and invent all the technologies needed for mobile computing in an open standard approach. This has resulted in the development of middleware API to support applications that need data replication and synchronization support, and a prototype application that demonstrates the functionality and advantages of the middleware.

Background and objectives: Gartner Group Inc. predicted that laptop computers would constitute 40% of all corporate PCs by 2000 and more than 70% of all major server/client applications will support disconnected operation. This is the basis for developing a middleware that provides general data synchronization support and helps enterprise application integration.

Technical discussion: For complicated, interactive applications, two-way synchronization and concurrent conflict reconciliation must be provided, along with increased scalability and fault-tolerance. Our middleware is superior to existing products in the market in these aspects: convergent data management and synchronization, building block based system construction, object-oriented conflict resolution, optimized data version management for one-to-one, client/server, and peer-to-peer systems, and efficient, adaptive and fault-tolerant data synchronization based on standard and popular web protocols.

Collaboration: This project has been fully funded by Network Computing Department (NC Dept.), Information Systems Laboratory (Jyohosouken), Mitsubishi Electric Corporation.

Future Directions: This research lasted for 3 years and has finished with an alpha version of the middleware software. Future work will be undertaken by the NC Dept. at Jyohosouken to improve the robustness of the preliminary alpha and beta releases of the software for inclusion in actual field trials by a well known wireless telecommunications company in Japan.

Contacts: Luosheng Peng, David Wong June 22, 2000

Schmoozer™ 3D Web Browser

See Color Plate #11

Background and objectives: Schmoozer is a 3D web authoring tool and run time environment. Schmoozer is a Java framework on top of the Open Community™ core. Users can build virtual worlds from existing models available on the web or from models they create. Then the worlds are connected together, making seamless transitions with Open Community's highly scalable architecture. Users also collaborate by using Schmoozer's multi-user editing and chat features.

Technical discussion: Open Community is a commercial version of the SPLINE (Scalable Platform for Large Interactive Networked Environments) architecture developed at MERL. Open Community implements ISTP (Interactive Sharing Transport Protocol) a protocol that improves scalability and allows distributed real time collaborative applications to be built over a network with no central server. This allows many users to share their virtual worlds just like HTTP browsers are used with web pages today.

The Schmoozer Java framework provides a user interface for navigating worlds and a drag and drop system for constructing or extending shared worlds. Schmoozer provides content developers with VRML2 support and XML scripting. Additional Java programming allows application developers to build sophisticated GUI interactions for collaboration and session management within Schmoozer. Schmoozer also supports Java's dynamic classes that make it is easy to add new object behaviors at run time. And Schmoozer provides an open XML format for saving and restoring worlds.

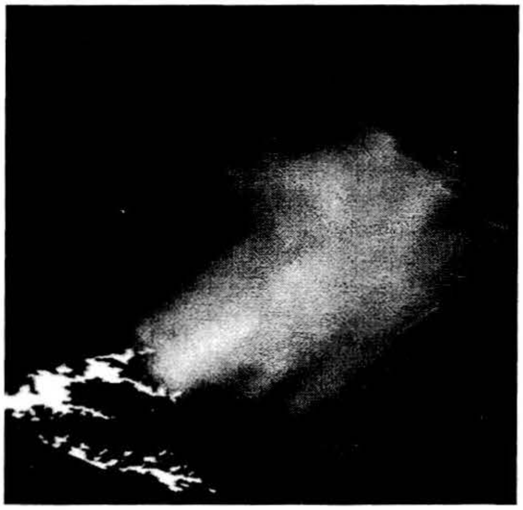
Schmoozer is currently available as a free download with install script at http://www.Schoozer.net. The download has already received interest from key developers of web technologies. Schmoozer has been used by the MIT Architecture department. Schmoozer was recently featured in the book "Networked Virtual Environments" by Singahal and Zyda from Addison Wesley. Schmoozer is currently waiting for good applications that can make use of it's 3D VR architecture.

Collaboration: Johosouken, OIT, MIT, MSS, RKC, Nearlife, Neometron, Draw, SGI.

Future Directions: Audio and Video improvements, Java Core, VRML X3D / MPEG standards.

Contacts: Derek Schwenke. May 26 2000

Weather Radar Data Visualization



The Weather Radar Group at Tsuden is developing weather sensing technologies that image the atmosphere at finer and finer resolutions. With such a massive amount of data being generated, the traditional visualization mechanisms used for the interpretation of weather radar data are insufficient. New visualization mechanisms are being developed to enable use of the information coming from these new weather sensing systems.

Background and objectives: Volume graphics visualization technologies from RTViz are being employed to visualize the weather data that is generated from Tsuden's weather radar products. The immediate objective of this project is one of creating 3D viewers that show the dynamics and physical extent of cloud structure in one unified, and accurately registered, display.

Technical discussion: The initial development of the weather radar visualization system will utilized the VolumePro 500 developed by RTViz (http://www.rtviz.com) for rendering the cloud and terrain data. Registration of the data will use a system, currently under development by MERL CAMBRIDGE SYSTEMS, that allows applications to register, resample, and reparameterize multidimensional sampled data. Development will take place under the Windows NT operating system.

Collaboration: The development of the Weather Radar Visualization System is a collaborative effort between MERL, Johosoken, and Tsuden. Tsuden will provide the domain expertise (i.e. expertise specific to weather radar), Johosoken will act as systems integrator, and MERL will provide basic technologies to Johosoken.

Future Directions: The development of the weather radar system will occur in two phases. The first phase will produce a tool that meteorological researchers will use to develop better weather prediction models utilizing volumetric data. The second phase will apply results in from the first phase in various transportation domains (e.g. aviation and railroads).

Contacts: Mike Walterman, Ned Muroi June 9, 2000

MERL Murray Hill

MERL Murray Hill is structured to address three major areas of research and development: Digital Video Technology, Digital Communications Technology and Digital Networks.

The Digital Video Technology (DVT) group is developing techniques for digital video compression, communication and multimedia applications. Their effort is directed towards the development of an improved HDTV receiver chipset, down-conversion decoding algorithms, video-coding algorithms and video description, indexing and retrieval applications. DVT is an active participant and contributor to the MPEG–4 and MPEG–7 standards.

The Digital Communications Technology (DCT) group focuses on signal generation and detection for multimedia wireline and wireless communication systems. Future consumer communication devices will be highly versatile, with voice, data and video services and access to the Internet. Efficient protocols are needed to ensure secure and high-quality services in diverse environments. The DCT group is formulating solutions to address these challenging problems.

The Advanced Digital Networks (ADN) group is focusing on the design and integration of Home Service Gateways linking external networks to home networks based on IEEE1394, Powerline, Wireless and Phoneline, and the development of data broadcasting applications. ADN is also exploring potential applications of MELCO's mobile networking technology to end-to-end solutions, and actively participating in and contributing to industry standards such as UPnP and HomePNA.

MERL Murray Hill has filed more than thirty patents arising from the above activities. Its expertise includes:

- **System knowledge** of digital video-coding algorithms, digital communications systems and algorithms, wireless and wireline communications, digital broadcasting and data-networking software, and multiple-standard modem technologies.
- **Industry standards** MPEG-2, MPEG-4, MPEG-7, XDSL, SDR, TIA-3GPP2 and CDMA; consumer electronics standards: ATSC, EIA CEA, ATTC and IEEE; home networking standards: HPNA, ATVEF and UpnP.
- **Fast prototyping** of real-time hardware and firmware, digital signal processing, data-acquisition techniques and use of field-programmable gate array (FPGA) devices.
- **System ASICs** design, bit-accurate modeling, VHDL, and evaluation of drivers, firmware and software systems.

MERL Murray Hill has developed close working relationships with four MELCO R&D labs.: the Imaging System Development Center (digital projection TV with 1394 interface), the Information Technology R&D Center (advanced digital broadcasting), the System LSI Business Center (second-generation and third-generation video decoder), and the Communication Systems R&D Center (4G wireless system). MERL Murray Hill and MELCO's Semiconductor Business Unit teamed with Lucent Technology on the design of the world's first digital HDTV receiver chip set. The laboratory is located in Murray Hill NJ, with easy access to New York City and major airports. The current staff consists of nineteen technical professionals, eight with PhD degrees and most of the remaining with MS degrees.

Technical Staff

Jay Bao

Ph.D., Senior Principal Member Technical Staff

Jay Bao received BS and an MS in Electronic Physics from Nankai University, Tianjin, China and a Ph.D. in Experimental Particle Physics from the Johns Hopkins University in 1992. From 1992 to 1995 he worked at Yale University as an associate research scientist, and joined MERL Murray Hill in August 1995. Interests include multimedia signal processing and digital communication systems.

Johnas Cukier

M.Sc., Principal Member Technical Staff

Johnas Cukier received his B.Sc. degree in Physics and Computer Science in 1983 from New York University and his M.Sc. degree in Electrical Engineering in 1985 from Polytechnic Institute of New York. He joined MERL Murray Hill in 1996, working on digital systems for CATV, RF microwave transmitters/receivers, and front-end advanced TV receivers. His current interests are in advanced TV, DSP, and RF transmitters/receivers.

Paul DaGraca

M.Sc., Principal Member Technical Staff

Paul Da Graca received a BS degree in Electrical Engineering from New Jersey Institute of Technology and a ME degree in Electrical Engineering from Stevens Institute of Technology. In 1996, he joined the MERL Murray Hill Laboratory. As a Principal Technical Staff member of MERL Murray Hill, he was a key member in the development of a HDTV Video Decoder IC for ATV receivers. Currently he is working on a 2nd generation HD/SD Video Decoder IC. His current research interests in the area of telecommunications include; demodulation techniques, digital signal processing, FPGA and ASIC design, as well as hardware and system design.

Parthapratim De

Ph.D., Member Technical Staff

Partha De received his Master's degree in Electrical Engineering from Indian Institute of Technology, Kanpur and his Ph.D. in Electrical and Computer Engineering from the University of Cincinnati in 1998, after which he joined MERL Murray Hill. His research interests are signal processing, image processing and communication systems.

Ajay Divakaran

Ph.D., Principal Member Technical Staff

Ajay Divakaran received his Ph.D. from Rensselaer Polytechnic Institute in 1993 and worked as a research associate at the Indian Institute of Science before joining Iterated Systems Inc., Atlanta, GA in 1995. At Iterated Systems he worked on video-coding algorithms for video telephony and entertainment-quality video. In 1998 he joined MERL Murray Hill, where he is working on video indexing with a view to MPEG-7 applications. His primary research interests are video indexing and coding.

James Fang

B. Sc., Member Technical Staff

James Fang received his B.Sc. from Columbia University in 1992 and did some graduate work there before joining Mitsubishi Electric in 1995. He worked on consumer televisions for three years before transferring to MERL Murray Hill in 1998. He is currently working on digital wireless communications.

Daqing Gu

Ph.D., Member of Technical Staff

Daqing Gu received the BE degree from Tsinghua University, Beijing, China in 1987; the MS and Ph.D. degrees in electrical engineering from the State University of New York at Stony Brook, Stony Brook, NY in 1996 and 1999, respectively. He is currently a Member of Technical Staff at MERL Murray Hill. His research interests include mobility management and radio resource management for wireless communications, wireless networking.

Jainlin Guo

Ph.D., Member Technical Staff

Jainlin Guo received his Ph.D. from Windsor University in 1995. He worked at Waterloo Maple for a year and a half as a software developer and then joined MERL Murray Hill in 1998. He has published seven research papers and his primary research interests include home networks, digital broadcasting, and wireless computing.

Jyhchau 'Henry' Horng

Ph.D. Member Technical Staff

Henry Horng received the Ph.D. from Polytechnic University in 1998. He has worked as a research assistant at Polytechnic and as software developer and lecturer for Chung Cheng Institute of Technology, Taiwan. Henry joined MERL Murray Hill in 1999. He has published seven research papers and has one patent. His primary research interests include digital signal processing and communications.

Fernando Matsubara

M.Sc., Senior Principal Member Technical Staff

Fernando Matsubara received his B.Sc. degree from the Universidad Nacional Autonoma de Mexico and his ME degree from the University of Tokyo. He managed mathematical simulation projects at the Institute for Electrical Research (Mexico) for five years and joined Mitsubishi Electric in 1990, where he managed projects on distributed computing and networking, moving to the U.S. in 1995. His primary research areas are digital communications, network technology, and digital video.

Robert Prozorov

M.Sc., Principal Member Technical Staff

Bob Prozorov received his M.S.E.E. from the Moscow Aircraft University in 1983 and then worked at the Computer Research Institute in Moscow. In 1990 he moved to General Electric and worked on design of power line/equipment protection systems. In 1996 he joined Lucent Technologies to participate in DSP code development for CDMA/AMPS digital cellular phones, and came to MERL Murray Hill in 1998, where his

Huifang Sun

Ph.D., Deputy Director Advanced Television Laboratory

Huifang Sun received his B.Sc. degree from Harbin Engineering Institute, Harbin, China in 1967, and his Ph.D. from University of Ottawa, Canada, in 1986. He was an Associate Professor at Fairleigh Dickinson University before moving to Sarnoff Research Laboratory in 1990, where he was Technology Leader of Digital Video Communication. He joined MERL Murray Hill in 1995. His research interests include digital video/image compression and digital communication.

Anthony Vetro

M.Sc., Principal Member Technical Staff

Anthony Vetro simultaneously received the BS and MS degrees in Electrical Engineering from Polytechnic University, Brooklyn, NY, in June 1996. He is currently pursuing his Ph.D. degree at the same university. He joined the MERL Murray Hill in April 1996 as a Member of the Technical Staff and in July of 1998 was promoted to a Principal Member. Upon joining, he worked on algorithms for down-conversion decoding, which has been implemented into MELCO's second-generation HDTV decoder chip. More recently, his work has focused on object-based video coding and transcoding, with emphasis on rate-distortion modeling and optimization techniques. He has published more than 20 journal and conference papers on these topics and has been issued two US patents with several more pending.

Charles You

M.Sc., Member Technical Staff

Charles You received the BS and MS in electrical engineering from the University of Buffalo, New York and the Columbia University, New York in 1995 and 1999, respectively. MERL Murray Hill has employed him since February 2000, working in the area of DSPs and wireless mobile systems. His main research interests are in the general areas of DSPs and communications.

Thierry Derand

M.Sc., Member Technical Staff

Thierry Derand is a student from a French college, E.S.I.E.E. (Ecole Supérieure d'Ingénieurs en Electrotechnique et Electronique), Paris, France. He will be graduated in July, 2000 with a French engineer diploma, equivalent to a MS desgree. He is doing a six-month internship within MERL Murray Hill, working on a project concerning the fourth generation of wireless communications. His main research interests are in general wireless communications.

George Carson

Ph.D., Consultant

George Carson received the Ph.D. from the University of California at Riverside in 1975. During his professional career he has worked in many diverse areas including computer-aided design, digital signal processing, automation, formal specification, modeling and simulation, computer graphics, image processing, computer communications and engineering management. His current research interests are in real time signal and information processing systems and in computer graphics and interactive techniques.

Major Recent Publications

Divakaran, Ajay., Peker, K., Sun, Huifang., "Descriptor for Spatial Distribution of Motion Activity for Compressed Video", *to appear Proc. IEEE Int'l Conference on Image Processing*, September 2000.

Lin, I. J., Divakaran, Ajay., Vetro, Anthony., Kung, S.Y. "Efficient Representation and Comparison of Multimedia Content Using DAG-composition," *to appear Proc. IEEE Int'l Conference and Multimedia Expo*, August 2000.

Divakaran, Ajay., Sekiguchi, S., Asai K., Sun, Huifang, "A Description Scheme For Video Based On Feature Extraction In The Compressed Domain," *Proc. Int'l Conference on Consumer Electronics*, June 2000.

Vetro, Anthony., Sun, Huifang., Y. Wang, "An MPEG-4 Object-Based Transcoder: Architecture And Applications," *Proc. Int'l Conference on Consumer Electronics*, June 2000.

Bao, Jay., "Optimization Of Transceiver Systems For Broadband Fixed Wireless Access Systems." *Proc. Int'l Conference on Consumer Electronics*, June 2000.

Vetro, Anthony., Sun, Huifang., Wang, Y., "Object-Based Transcoding For Scalable Quality Of Service," *Proc. Int'l Symposium on Circuits and Systems*, May 2000.

Divakaran, Ajay., Ito, H., Sun, Huifang., Poon, Tommy., "Fade-In/Out Scene Change Detection In MPEG-1/2/4 Compressed Domain," *Proc SPIE Conference on Storage and Retrieval for Media Databases*, January 2000.

Divakaran, Ajay., Sun, Huifang., "A Descriptor For Spatial Distribution Of Motion Activity" *Proc. SPIE Conference on Storage and Retrieval for Media Databases*, January 2000.

Bao, Jay., Zeng, S., You, Charles.,"A Programmable DSP-Based Approach To The Front-End Of Demodulators For Digital TV Receivers", *Proc. International Conference on Signal Processing Application and Technology*, November 1999.

Lin, I-J., Vetro, Anthony., Divakaran, Ajay., Kung, S-Y. "Shape Representation And Comparison Using Voronoi Order Skeletons And DAG Ordered Trees," *Proc. Int'l Workshop on Very Low Bitrate Video Coding,* Oct. 1999.

Vetro, Anthony., Sun, Huifang., Wang, Y., Guleryuz, O., "Rate-Distortion Modeling Of Binary Shape Using State Partitioning," *Proc. Int'l Conference on Image Processing*, Oct. 1999.

Duardo, O., DaGraca, Paul., Hosotani, S., Sugawa S., Jiang, H., ``A Cost Effective HDTV Decoder IC With Integrated Graphics Engine And Display Processor," *IEEE Trans. Consumer Electronics*, vol.45, num.3, August 1999.

De, Parthapratim., Bao, Jay., Poon, Tommy., ``A Calculation-Efficient Algorithm For Decision Feedback Equalizers," *IEEE Trans. Consumer Electronics*, vol.45, num.3, August 1999.

De, Parthapratim., Bao, Jay., Poon, Tommy., " Subspace Based Decision Feedback Equalization", *Proc. IEEE Radio & Wireless Conference*, August 1999.

DaGraca, Paul., Duardo, O., Hosotani, S., Sugawa, S., Jiang, H. ''A Cost Effective MPEG-2 HDTV System And Video Decoder IC With Integrated Down-Decoder, Graphics And Display Processor," *Proc. Int'l Conference on Consumer Electronics*, June 1999.

De, Parthapratim., Bao, Jay., Poon, Tommy., ''A Calculation-Efficient Algorithm For Decision Feedback Equalizers," *Proc. Int'l Conference on Consumer Electronics,* June 1999.

Poon, Tommy., ''Key Enabling Technology For HDTV," *Proc. IEEE Symposium on Advances in Wired and Wireless Communications*, March 1999.

Vetro, Anthony., Sun, Huifang., Wang,Y., ''MPEG-4 Rate Control For Coding Multiple Video Objects," *IEEE Trans. Circuits Syst. Video Technology*, vol. 9, no. 1, pp. 186-199, February 1999.

Divakaran, Ajay., Ito, H., Sun, Huifang., and Poon, Tommy., ''Scene Change Detection And Feature Extraction For MPEG-4 Sequences," *Proc. SPIE Conference on Storage and Retrevial for Image and Video Databases*, January 1999.

Vetro, Anthony., Sun, Huifang., Chen Y.K., and Kung, S.Y., ''True Motion Vectors For Robust Video Transmission," *Proc. SPIE Conference on Visual Communications and Image Processing (VCIP '99)*, January 1999.

Chen, Y.K., Vetro, Anthony., Sun, Huifang., and Kung, S.Y., ''Frame-Rate Up-Conversion Using Transmitted True Motion," *Proc. Second Int'l Workshop on Multimedia Signal Processing (MMSP '98)*, Dec. 1998.

Bao, Jay., Lu, C.Y., DaGraca, Paul., Zeng, S., Poon, Tommy., "A New Timing Recovery Method for DTV Receivers", *IEEE Trans. On Consumer Electronics* , Vol.44, No 4, November 1998.

Bao, Jay., "A Reconfigurable Adaptive Equalizer for ATV Receivers", *Proc. Int'l. Conf. on Signal Processing Appl. & Technology* (ICSPAT 98), September 1998.

Vetro, Anthony., Sun, Huifang., DaGraca, Paul., Poon, Tommy., ''Minimum Drift Architectures For Three-Layer Scalable DTV Decoding," *IEEE Trans. Consumer Electronics*, vol.44, num.3, August 1998

Tsunashima, K., Muneishi, K., Morikawa, H., Kayashima, H., Sinyanskiy, V.,''An integrated DTV receiver for ATSC digital television standard," *IEEE Trans. Consumer Electronics*, vol.44, num.3, August 1998.

Sinyanskiy, V., Cukier, Johnas., Davidson, A., Poon, Tommy., ''Front-End Of A Digital ATV Receiver," *IEEE Trans. Consumer Electronics*, vol.44, num.3, August 1998.

Vetro, Anthony., Sun, Huifang.,''Frequency Domain Down-Conversion Using An Optimal Motion Compensation Scheme," *Int'l Journal of Imaging Systems & Technology*, vol.9, num.4, August 1998.

Vetro, Anthony., Sun, Huifang., ''On The Motion Compensation Within A Down-Conversion Decoder," *Journal of Electronic Imaging*, vol.7, num.3, July 1998.

Vetro, Anthony., Sun, Huifang., Wang, Y., ''Joint Shape And Texture Rate Control For MPEG-4 Encoders" *Proc. Int'l Symposium on Circuits and Systems (ISCAS '98)*, June 1998.

Vetro, Anthony., Sun, Huifang., DaGraca, Paul., Poon, Tommy., ''A Minimum Drift Three-Layer Scalable DTV Decoder," *Proc. Int'l Conf. Consumer Electronics (ICCE '98)*, June 1998.

Tsunashima, K., Muneishi, K., Morikawa, H., Kayashima H., Sinyanskiy, V., ''An Integrated DTV Receiver For ATSC Digital Television Standard," *Proc. Int'l Conf. Consumer Electronics (ICCE '98)*, June 1998.

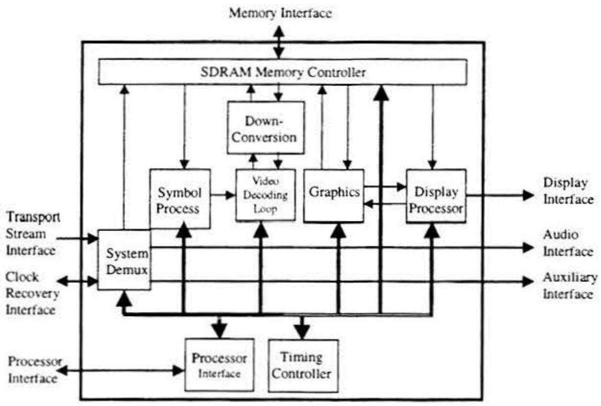
Sinyanskiy, V., Cukier, Johnas., Davidson A., Poon, Tommy., ''Front-End Of A Digital ATV Receiver," *Proc. Int'l Conf. Consumer Electronics (ICCE '98)*, June 1998.

Bao, Jay., Lu, C.Y. , DaGraca, Paul., Zeng, S., Poon, Tommy.," ''A New Timing Recovery Method For DTV Receivers," *Proc. Int'l Conf. Consumer Electronics* , June 1998.

Vetro, Anthony., Sun, Huifang., ''Generalized Motion Compensation For Drift Minimization," *Proc SPIE Conference on Visual Communications and Image Processing (VCIP '98)*, January 1998.

Project Descriptions

A 2nd Generation Video Decoder IC



A highly integrated and cost effective second generation HDTV IC that supports MPEG-2, ATSC, BS4 and other standards is being developed for SD TV applications.

The IC performs system and video decoding. It incorporates MERL Murray Hill's Down Conversion algorithm for viewing High definition broadcasts on 480I/P displays. Video data is output using built-in display processing and 2D graphics functions. Other features include audio and auxiliary data outputs, VCR interface, audio and video clock recovery, and supporting DNTSC input.

HDTV: High Definition Television, IC : Integrated Circuit, ATSC: Advanced Television System Committee, DNTSC: Digitized Nation Television System Committee, SD: Standard Definition.

Background and objectives: Over fifty television stations are currently broadcasting digital television signals in the US. In order for all consumers to benefit from features that are being broadcast in the ATSC format, a highly integrated and cost effective video decoder IC that can convert from HD to SD format for viewing on today's NTSC type televisions must be developed.

Key Design Features: This IC is very cost effective because it is able to optimize external memory storage. It must handle the storing of audio, compressed, reference picture, SI tables, and graphics data into low cost 16Mbit SDRAM. This accomplishment is made possible due to the integrated down conversion logic (developed at MERL Murray Hill) that re-scales HD input pictures to a quarter of there size so that only three SDRAM are required for storage.

The IC includes a programmable MPEG2 System Demultiplexer that supports transport streams up to 135 Mbps, video and audio clock recovery, synchronization, and 32 independent SI filters and queues. The video decoding function supports automatic error concealment. The internal display processor outputs raster data for 480I/P displays by performing 34 different video format conversions. Finally, the IC contains full-featured 2D graphics functions controlled via a processor interface.

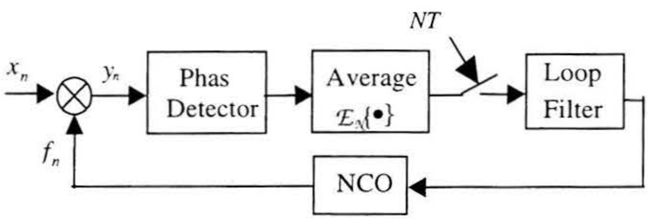
Collaboration: This project is being developed jointly by MERL Murray Hill, New Providence, New Jersey, Mitsubishi Electric – System LSI Development Center (L-Ji-Se) Mizuhara, Itami, Hyogo, Japan, and also with help from Bell Labs, Lucent Technologies, Murray Hill, NJ.

Future Directions: A potential next generation of this IC would incorporate internal or embedded SDRAM memories, an audio decoder and full featured display processor.

Contact: Paul Da Graca

April 11, 2000

A Programmable DSP-based PLL for Digital TV Receivers



A programmable DSP-based phase locked loop (PLL) for DTV receiver carrier and timing recovery loops was developed. Instead of processing the incoming data at sampling rate, A block of data with N samples are processed at a time. The block sizes can be varied based on the signal condition, and the state of the phase locked loop. This results in more flexible front-end design, which is an important factor for multi-mode receivers. Three different implementations of the block-based approach are described, together with simulation results and DSP implementation.

CR: Carrier Recovery, *PLL*: Phase Locked Loop, *NCO*: Numerical Control Oscillator

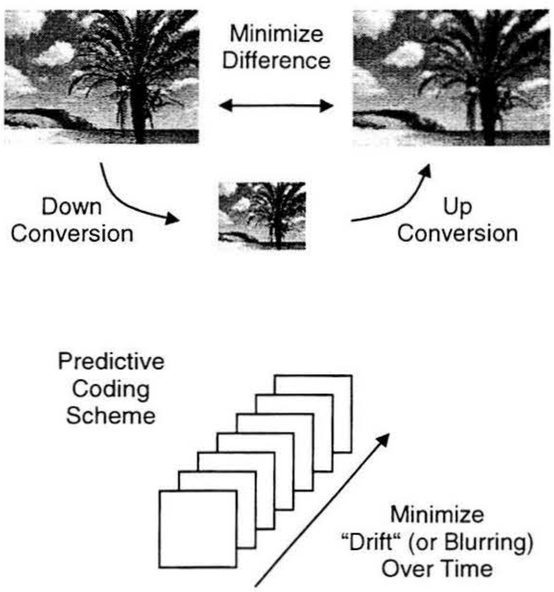
Background and objectives: The era of digital television broadcasting in the U.S. began with the introduction of terrestrial services in November 1998. Cable and satellite digital TV broadcasting services will soon become available as well. Because of the high similarity in source coding used throughout all three transmission media (terrestrial, cable and satellite), where MPEG-2 is the common standard for video and audio coding, it is possible to share signal processing functional blocks across the three media. In current generation digital TV receivers, due to the high data throughput, (e.g. the symbol rates: 10.76MHz for 8VSB, 5.38MHz for 256QAM, and up to 45MHz for QPSK,) the conventional approach of implementation is hardware. The hardware approach can offer a high computation speed but is very difficult to change for future upgrade, and the size of the chip is usually quite large, typically accompanied with a long development cycle. In this paper, we present a new solution to the multi-mode demodulator front-end that is based on programmable DSP, and re-configurable logic.

Technical discussion: The function of phase detector block is to extract phase error from equalizer output x_n. An integrator for average is introduced prior to the loop filter. This average block takes a block of N data from the phase detector and put the average to the loop filter every block, *Ts* is the sampling interval. Since the number of operation is reduced for the loop filter and NCO, both parts can be implemented in DSP software. The block size N can vary at the different stage, such as the acquisition and tracking state.

Future Directions: This algorithm can be applied to wireless mobile systems.

Contacts: Jay Bao, Charles You March 30, 2000

Down-Conversion Filters for Improved Picture Quality



Down-conversion is a useful technology to reduce memory requirements and complexity within DTV receivers. They also enable the use of eRAM technology, a strong point for MELCO. The algorithms for down-conversion have been extensively studied and evaluated at MERL Murray Hill. The filters to perform down-conversion, along with the optimal up-conversion filters associated with these filters, have been implemented into an integrated Bit Accurate Model (BAM) of an HDTV Receiver. The receiver includes Demux, Decode and Display Processing components. This BAM was used to verify the operation of the MELCO's 2nd generation HDTV receiver LSI. Further work to improve the quality produced by the down/up-conversion process is being investigated.

Background and objectives: The project was initiated in 1995 to conduct research on new algorithms for memory reduction within a video decoder. Not only are memory requirements reduced, but implementation costs are also reduced due to impact on memory bandwidth and clock rate. Two key technologies have been the major focus: down-conversion and motion compensation. These new algorithms are basic components of a low-cost video decoder and may be used in a wide range of image/video products.

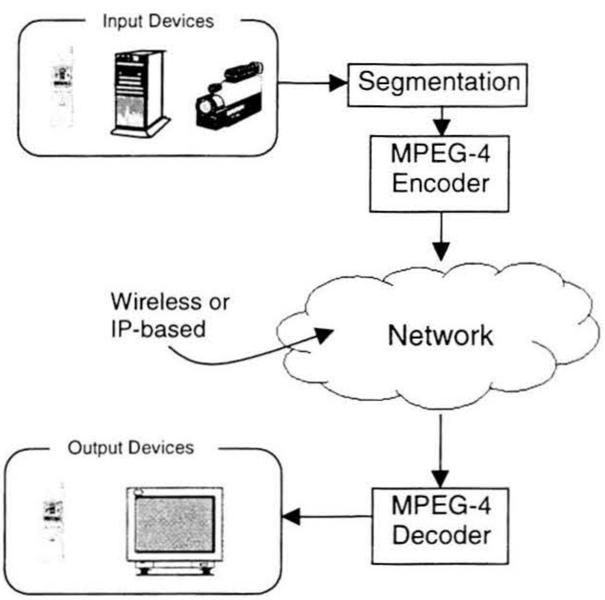
Technical discussion: The filters used for down-conversion are based on the concept of frequency synthesis, and the filters used to perform the up-conversion are determined by an optimal least-squares solution. The combination of these techniques allows us to achieve significant reductions in the amount of observable drift compared to previously published methods.

Collaboration: The initial funding for this project came from Eijo-ken. Several members of this lab provided valuable input on a regular basis. This input was useful in shaping the direction of the project and evaluating our progress. In 1998, collaboration with L-Ji-Se became stronger as several Japanese engineers were assigned to the development of the 2nd generation HDTV receiver LSI, which implemented the down-conversion algorithms. Finally, Johosoken has provided expert evaluation regarding the quality of the down-decoded video sequences.

Future Directions: Our current effort is to improve the quality that can be achieved in the down/up-conversion process. Emphasis is being place on improved line quality, overall sharpness and drift. Additionally, any impacts on the architectural design are being considered.

Contacts: Anthony Vetro Huifang Sun April 1, 2000

Encoding Algorithms for MPEG-4 Systems and Applications



The MPEG-4 standard has just been completed last year. It is the first standard to support object-based coding. Also, error resilience is a key feature of this new standard, making it suitable for video transmission over error-prone channels. The first application of MPEG-4 will be mobile videophone terminals. As part of this project, we are developing algorithms for efficient object-based encoding and transmission. On the encoder side, this includes object segmentation algorithms, rate control for multiple objects and fast motion estimation. To assist with transmission, we consider efficient means of transcoding for bit-rate reduction and reduced resolution.

Background and objectives: The goal of this project is to develop algorithms that support an MPEG-4 object-based coding and delivery system. We have focused our research efforts on critical parts of the encoder that affect the quality and complexity, namely the rate control and fast motion estimation. In order to support object-based functionalities, we have also considered algorithms to perform the object segmentation. Finally, to assist in the delivery of MPEG-4 bitstreams over a network that is bit-rate constrained or to a user device that is limited in display size or computational power, we have considered algorithms to perform transcoding. The purpose of the transcoder is to convert the original bitstream into a new bitstream that meets constrains imposed by the network or limitations in the user device.

Technical discussion: A rate control scheme to support the object-based functionality of MPEG-4 has been developed. The rate control for multiple video objects is unique in that shape information must also be coded along with texture and motion for each object. For object segmentation, the algorithm can be broken down into several steps. The first step estimates an initial object boundary for a sequence of images. The second step orders the space exterior to each image for effective searching. The final step minimizes an energy function that incorporates visual discontinuity, motion discontinuity and smoothness over time. To perform object-based transcoding for bit-rate reduction, an efficient dynamic programming approach has been developed. This algorithm is capable of making optimal trade-offs in spatial vs temporal quality.

Collaboration: This project is done in collaboration with the Multimedia Information Coding & Transmission Technology Department at Johosoken and Princeton University.

Future Directions: The focus of future work will be on systems and applications that integrate this technology with multimedia search and retrieval. Consequently, we plan to consider a real-time multimedia client-server environment.

Contacts: Anthony Vetro and Huifang Sun March 31, 2000

Low Complexity Adaptive Equalization and Multiuser Detection

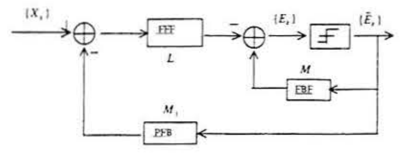
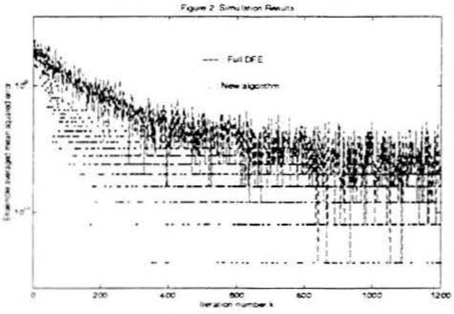


Figure 2 Simulation Results



High speed data transmission is a characteristic of digital television broadcasting. Future wireless systems will also handle broadband data. This requires sophisticated receiver structures in high mobility environment. Mobile multimedia access from handheld devices is one such example. We introduce new algorithms for low complexity adaptive equalization and multi-user detection. They can be applied to digital receivers designed for wideband multimedia applications, such as WCDMA and future generation wireless terminals.

WCDMA: Wideband Code Division Multiple Access

Background and objectives: Efficient interference cancellation techniques are developed for high speed data transmission. This includes digital television broadcasting with data rates of 12 Mbits/s or more. Similarly, third and fourth generation wireless systems are expected to handle very high data rates.

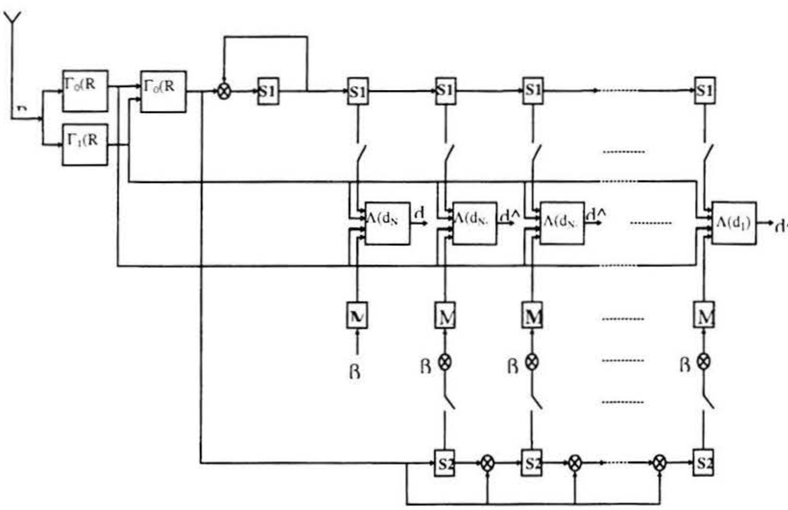
Technical discussion: Due to propagation conditions, multipath signals hamper the reception of the main signal. An equalizer has to be designed to compensate for this effect. For very high data rates, the multipath delays might be appreciable. In this project, we developed robust, calculation efficient adaptive equalizer algorithms using sparse channel characteristics of terrestrial transmission, together with a partial feedback structure, as illustrated above. This results in about ten times less calculations and reduced hardware complexity, and about 50 percent reduction in decision errors. The new algorithm also has fast response to dynamic channel impairments. In mobile cellular systems, different users share a common channel by modulating a set of signature waveforms. The receiver gets the sum of transmitted signals and noise. Subspace based algorithms can be used to obtain more efficient detectors and better performance. These detectors can be used in wireless multimedia systems. It can identify and demodulate multiple users at the base station, or identify and demodulate selected user at the handset.

Collaboration: As a part of the strategic advanced project on wireless communications, we work with ITC and ITE researchers.

Future Directions: We will continue to work on system level design and characterization for WCDMA environment.

Contact: Parthapratim De   March 30, 2000

Matrix-based Turbo Decoder Architecture for 3G Wireless Systems



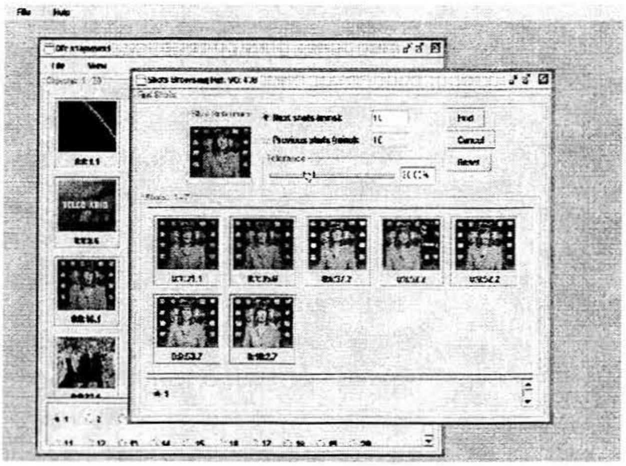
We introduce a matrix-based turbo decoder architecture that explores the parallelism of matrix calculations. By formulating calculation of key transition probabilities into matrix forms, and by shortening matrix multiplication time down to one column by one row multiplication and summation, a turbo decoder with up to 4 times faster compared to conventional approaches. The architecture is also suitable for efficient VLSI implementation.

Background and objectives: In mobile wireless communications, intersymbol interference due to fading and multipath impairments causes error in received signals. Thermal noise induced by electronic components also causes errors. In recent years, parallel concatenated recursive convolutional codes, i.e. turbo codes, has shown great performance gain over conventional error correction codes, and thus has become a part of the system standards for a number of 3^{rd} generation wireless system standards, such as cdma2000. Various turbo decoder algorithms have been invented, and MAP (Maximum *a posteriori*) method has proven to be a superior iterative decoding algorithm. However, conventional MAP is complex to implement due to computation complexity of forward-backward recursions required, and memory demand. It is therefore important to find efficient and low complexity algorithms for implementing MAP in VLSI.
Technical discussion: We introduce a matrix transform based method for MAP decoding of turbo codes. The successive decoding procedures are performed in parallel and well formulated into a set of simple and regular matrix operations. These operations substantially accelerate decoding and reduce the computational complexity, and are particularly suited for implementation in special-purpose parallel processing VLSI hardware architectures. Our method begins by initializing a forward recursion probability function vector α_0, and a backward recursion probability function vector β_N. We then determine transition probability matrices $\Gamma(R_k)$ and $\Gamma_i(R_k)$ for each received symbol R_1^N. From $\Gamma(R_k)$, we then determine values of α_k which corresponds to received symbol R_k. Meanwhile, multiple multiplications on $\Gamma(R_k)$ and $\Gamma_i(R_k)$ are done in parallel. After receiving the complete symbol sequence R_1^N, values of backward recursion probability β_k are calculated in parallel, and thus the log likelihood ratio for each decoded bit d_k. As shown in accompanying figure, a shift register based implementation effectively reduces memory requirements and simplifies data access.
Collaboration: Dr. D. Wang and Professor H. Kobayashi of Princeton University.

Future Directions: Apply to 3G receiver designs, explore new turbo decoding architectures based on believe propagation (with Dr. W. Freeman of MERL Cambridge Research).

Contacts: Jay Bao June 26, 2000

MPEG-7 Compliant Video Browsing and Indexing System



See Color Plate #12

As more and more audio-visual content becomes available in digital form, the ability to locate desired content will become more and more important. The emerging MPEG-7 standard will standardize a multimedia content description interface that will enable efficient searching and browsing of worldwide multimedia content. At the MERL Murray Hill we are developing video indexing techniques to enable rapid browsing and querying of local and remote video databases. We have proposed a video descriptor to MPEG-7, which is currently in the working draft of the standard. The screen shot at the left illustrates our content based video retrieval system.

Background and objectives: The proposed MPEG-7 standard provides a framework for browsing and querying of video content. Browsing and querying of video content requires effective video indexing techniques that rely on feature extraction. Since video content is increasingly available in digital compressed formats, feature extraction that operates directly in the compressed content is useful. Such extraction is fast because it does not incur the expense of full decoding and capitalizes on the information already computed by the original video compressor. Therefore, our focus is on feature extraction from compressed video. The main objective at this stage is to produce a video indexing system that allows real-time interactive retrieval using the descriptor and description scheme technology provided by MPEG-7.

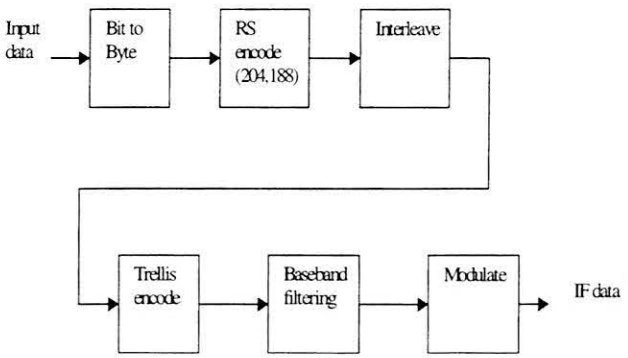
Technical discussion: One of the biggest challenges in feature extraction from the compressed domain is the vulnerability to changes in the encoding parameters. Motion vectors and intensity and color information are more or less robust to such changes. In our work we focus on extracting the intensity of motion activity in a video shot as well as the spatial distribution of motion activity in a frame, using the motion vectors from the compressed video bitstream. We use a simple run-length based representation for the spatial distribution. Our feature extraction is fast and effective and has been tested extensively within a prototype video browsing system.

Collaboration: Collaborators include the Multimedia Information Coding & Transmission Technology Department at Johosoken and the Visual Information Lab (VIL) of ITE.

Future Directions: We plan to add more descriptors to our video-browsing system and intend to continue our work on feature extraction in the compressed domain.

Contacts: Ajay Divakaran and Huifang Sun March 31, 2000

Optimization of Transceiver Design for Broadband Wireless Access



Radio Systems designed for broadband wireless multimedia services must combat fading and multipath impairments. We analyzed the effects of these impairments on transceiver signal processing system design, and propose a system architecture of spatial transmit diversity combined with adaptive modulation. The performance of such a system is compared favorably against conventional single carrier fixed modulation and OFDM systems.

OFDM: Orthogonal Frequency Division Multiplexing, MMDS: Microwave Multi-point Distribution Services

Background and objectives: Recently, broadband fixed wireless access technology has been widely regarded as an effective way to realize point-to-multipoint high data rate, multimedia communications. Broadband wireless access systems for multimedia applications must combat against fading, static and dynamic multipath impairments. The very high data rates (20Mb/s or higher) at low packet error rate for multimedia data, together with high carrier frequencies, ranging from time-dispersive MMDS bands (<10GHz), to fading dominated millimeter-wave bands (e.g. 10-66GHz), make these challenges particularly difficult.

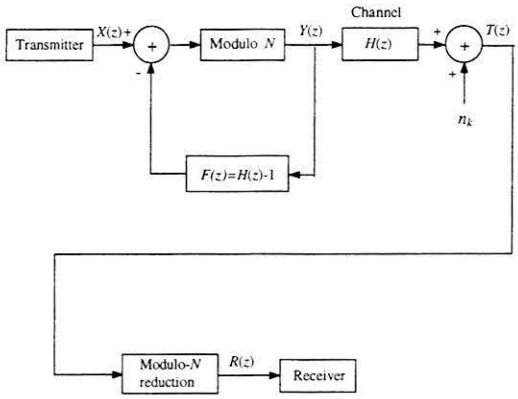
Technical discussion: In this work, optimal signal processing design for the transceiver systems is analyzed by combining spatial diversity and adaptive modulation. Spatial diversity is achieved using transmission diversity and adaptive antenna system, which improve throughputs and multipath interference rejection; adaptive modulation increases spectral efficiency and coverage by allowing different levels of services to users, based on use cases and channel link quality. The BER performance of the proposed system was examined using typical propagation models. Both indoor and outdoor propagation lose were modeled. A two-antenna transmit diversity system was used. Improvements in effective service coverage, and increase in throughputs for fixed service coverage were measured. It was found that more than 4.5dB diversity gain is achievable with the system described here. Additional gains can be achieved by also applying spatial diversity to reverse link from mobile terminals.

Collaboration: At MERL Murray Hill collaborates with ITC and ITE in a common SAP project for 4th generation wireless communication R&D. Our effort is also partially funded by Tsukaise.

Future Directions: Combine with precoding techniques to further improve performance.

Contact: Jay Bao March 29, 2000

Tomlinson-Harashima Precoder for Broadband Wireless System



In digital wireless communications, intersymbol interference (ISI) often occurs due to non-ideal pulse shaping and time-dispersive channels. Equalization techniques are usually used to compensate for ISI. The goal of this project is to simulate Tomlinson-Harashima modulo-type equalizer, and compare it with the other equalization techniques. The simulation results are going to be used to evaluate the equalization techniques for the next-generation broadband wireless communication systems.

ISI: Intersymbol Interference; AWGN: Additive White Gaussian Noise.

Background and objectives: We search for a suitable equalization and coded transmission technique for the next-generation wireless communication systems. The primary objective of this project is to build a simulation model for Tomlinson-Harashima precoder, and simulate it for a variety of channel and noise parameters.

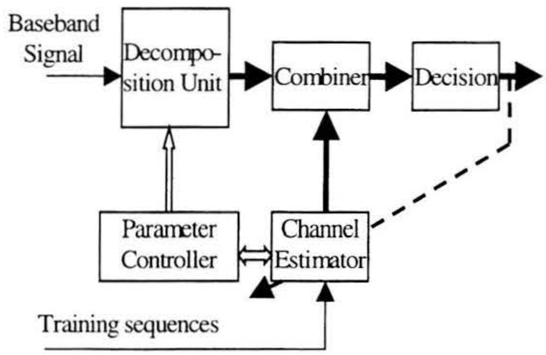
Technical discussion: In conventional equalization techniques, an equalizer would take the form of feedback transversal filters, which require an infinite or a very large number of stages. Otherwise, the compensation for ISI is only approximate. This large number of feedback stages will cause the error propagation problem. By using a Tomlinson-Harashima precoder at the transmitter end, these problems can be resolved. As illustrated in the figure above, the data coming from transmitter are pre-equalized with modulo-*N* inverse filter. The output obtained from the pre-equalizer is transmitted to the channel. It is shown that the overall system transfer function is unit if the data after transmission is modulo-*N* reduced. The modulo-*N* adder at the pre-equalizer is to eliminate the possibility of instability of the circuit. An AWGN noise source is added to the channel for the purpose of testing, simulating a physical channel.

Collaboration: As a part of R&D program, we intend to collaborate with Johosoken and ITE for the development of the next-generation wireless communication systems.

Future Directions: As a part of development of the next-generation wireless communication systems, we are currently investigating equalization algorithms for improvement over the currently used methods. In near future, a modulo-type Viterbi decoder in conjunction with Tomlinson-Harashima precoder will be investigated.

Contacts: Daqing Gu and Jay Bao March 30, 2000

Reconfigurable Digital Receiver for Multiple Multiplexing Schemes



With the existence of many industry standards, there will be great demand for a digital receiver that is capable of operating in multiple modes and multiple functions. This digital receiver has general hardware architecture with software features. By only changing the system parameters, it can be reconfigured in real time to adapt to different multiplexing schemes, such as CDMA, TDMA, FDMA, and OFDM. A user can operate the receiver in different geographical areas, or select a particular communication service provider at will. This receiver also uses adaptive time-frequency diversity combining techniques to improve the receiver performance over high mobility, high data rate communication channels.

CDMA: Code Division Multiple Access, TDMA: Time Division Multiple Access, FDMA: Frequency Division Multiple Access, OFDM: Orthogonal Frequency Division Multiplexing

Background and objectives: The explosive growth in telecommunication and multimedia applications demands flexible, efficient, high performance receivers. Most current digital receivers focus on detecting transmitted signals for a particular modulation or multiplexing scheme. With the existence of many different communication standards, flexible receiver architecture is necessary for next generation communication systems. The primary objective of this digital receiver is to provide reliable communication performance over high mobility, high data rate channels for different communication standards.

Technical discussion: Diversity techniques are used in practice to combat fading channels. The increased mobility of users in cellular communication often results in fast fading or Doppler spreads. The use of joint time-frequency diversity techniques provides significant performance improvement over the existing systems. Therefore, a time-frequency-diversity based receiver architecture with software features becomes a potential candidate of software defined radio for next generation wireless communication systems. It is noted that, unlike the time-frequency diversity is the wireless channel characteristics, one can create diversity dimension artificially, such as space-time diversity. Diversity gains can be achieved by using multiple elements for antenna arrays processing. The potential application of space-time processing is the multi-user detection in CDMA systems. However, due to heavy computational complexity, the trade off between performance and complexity become an importance issue when applying this technique.

Collaboration: MERL Murray Hill collaborates with Princeton University for the development of diversity combining techniques for future wireless communication systems.

Future Directions: Combine space-time-frequency diversity and reduce the complexity.

Contact: Henry Horng March 29, 2000

COLOR PLATES

Plate 1 – See Page 30

Plate 2 – See Page 33

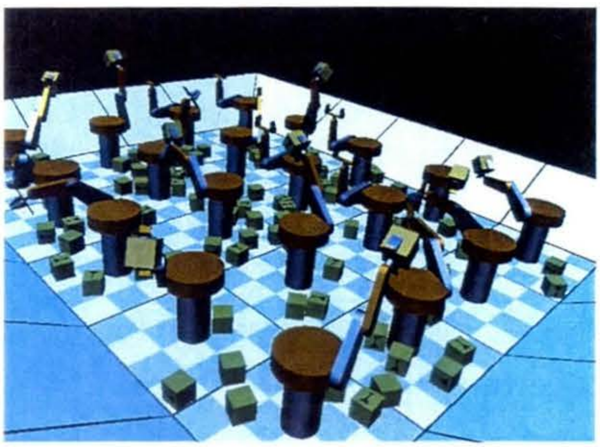


Plate 3 – See Page 35

Plate 4 – See Page 36

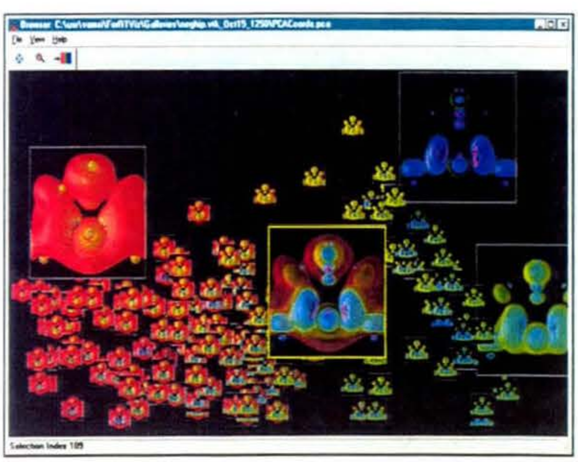


Plate 5 – See Page 37

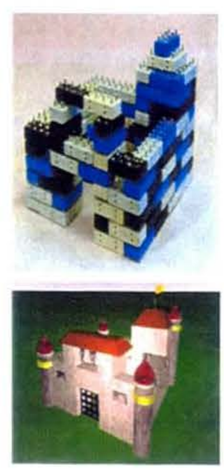


Plate 6 – See Page 41

COLOR PLATES

Plate 7 – See page 44

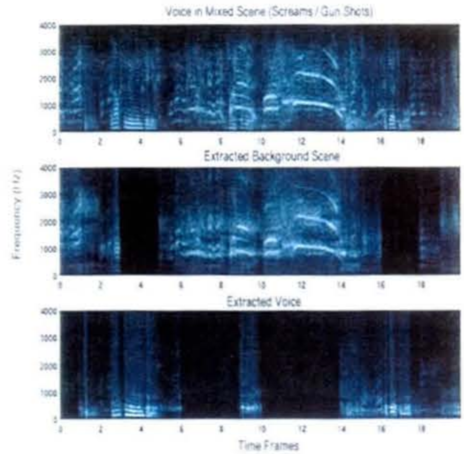


Plate 8 – see page 51

Plate 9 – See Page 57

Plate 10 – See Page 64

Plate 11 – See Page 81

Plate 12 – See Page 98