Mitsubishi Electric Information Technology Center America (ITA)

ITA Annual Report

July 1998 through June 1999

Welcome to ITA – Mitsubishi Electric Information Technology Center America, the North American corporate R&D arm of Mitsubishi Electric Corporation (MELCO). In this report you will find descriptions of ITA as a whole, and of our three laboratories, MERL – Mitsubishi Electric Research Laboratory, HSL – Horizon Systems Laboratory, and ATL – Advanced TV Laboratory. We also describe VGO – Volume Graphics Organization, which is in the process of moving into a separate business unit.

201 Broadway, Cambridge, Massachusetts 02139
617.621.7500

Production:

John Howard, Ron Johnson, Janet O'Halloran, Richard Waters

Table of Contents

Letter from the Chairman

It is my privilege to introduce you to Mitsubishi Electric Information Technology Center America (which we shorten to ITA), the North American Corporate R&D arm of Mitsubishi Electric Company (MELCO).

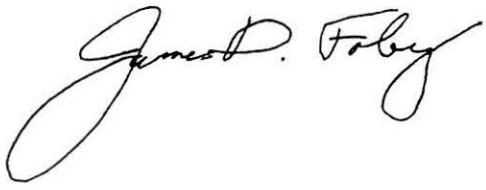
ITA conducts application-motivated basic research and advanced development in computer and communications technology. We seek to contribute to the advancement of science and to meet human needs by solving fundamental problems that arise in expanding the productive use of computers. Our vision is both long-term and market-driven, so we seek to anticipate and meet market needs and business opportunities. This means exploring entirely new categories of possibility 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 research results to contribute to MELCO's business success.

ITA focuses on three key technology sectors: advanced graphics and imaging technologies, internet computing and applications, and digital broadcasting communications. Each of our labs "owns" specific segments of these technologies, while working collaboratively to achieve functional, deliverable results. Our output ranges from application-motivated research in basic computer science, as represented by papers, patents, and proof-of-concept hardware and software prototypes, through industry-first products.

ITA is small enough to be agile and flexible in the dynamic marketplace of ideas and innovations, yet we can leverage the size, recognition and diversity of our strong global parent. We turn our technical achievements into business successes by partnering with business units and with other labs in MELCO's global corporate R&D network. Our efficient size and resulting flexibility encourages each individual laboratory to partner with both internal and external organizations to meet their objectives through shared innovation and knowledge. In building partnerships, we shape our technologies to meet real market opportunities related to MELCO businesses. With selected projects, we do early business development directly with customers, while also working with MELCO or other partners who assume final business responsibility.

Our location in North America lets us perform our R&D in what is currently the world's liveliest hotbed of technological innovation and business opportunities. We are involved in the R&D community and standards activities, we are aware of market trends, and we interact with local MELCO customers and business partners. We also maintain long-standing cooperative relationships with research universities such as Brown, CMU, Columbia, Georgia Tech, MIT, 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.meitca.com".



James D. Foley
Chairman and CEO

ITA – Mitsubishi Electric Information Technology Center America

Welcome to ITA – Mitsubishi Electric Information Technology Center America, the North American corporate R&D arm of Mitsubishi Electric Corporation (MELCO). In this report you will find descriptions of ITA as a whole, and of our three laboratories, MERL – Mitsubishi Electric Research Laboratory, HSL – Horizon Systems Laboratory, and ATL – Advanced TV Laboratory. We also describe VGO – Volume Graphics Organization, which is in the process of moving into a separate business unit.

- **ATL** – Advanced Television Laboratory, in New Providence, NJ, specializes in digital TV – the convergence point of television and computer technologies – as well as digital communications. ATL works closely with MELCO'S semiconductor and audiovisual business units and with MELCO'S Consumer Electronics Engineering Center in California. ATL also facilitates MELCO's partnership with Lucent for digital TV chip set development: the first chip set is used in MELCO's and others' HDTV products. ATL's industry-leading HDTV down-conversion algorithms are being incorporated into a second-generation chip set, giving MELCO an important competitive advantage.
- **HSL** – Horizon Systems Lab in Waltham, MA, is developing several innovative Internet software applications and systems: Open Community – for distributed collaboration; Concordia – Java mobile agent software; NetRep – network file management (jointly with Veritas); Wicked Mail – a Java IMAP-4 Email client; and home networking. Concordia is being used by many groups within MELCO and has been licensed to several companies in North America. NetRep is generating license revenues and in a recent key competitive test outperformed its competition by a factor of three.
- **MERL** – Mitsubishi Electric Research Lab, in Cambridge, MA, is our basic computer science research lab, with diverse research specialties ranging from real-time distributed systems and networking through computer graphics, computer vision, and expert systems to human-computer interaction. MERL's MidART network middleware is part of a shipping product, and vision research has helped create business opportunities for MELCO chips. MERL graphics research created and continues to support our volume graphics product development activities.
- **VGO** – Volume Graphics Organization, also in Cambridge, MA has designed the first ASIC chip capable of real-time volume rendering on a PC plug-in board. VGO provides advanced development and engineering support to Real Time Visualization, the new MELCO business entity responsible for volume graphics. First shipments of this product were made in May 1999. The medical and geophysical industries are targeted as early users of volume graphics. A second-generation chip with additional features and enhanced performance is under development for shipment next year.

Within ITA, we emphasize the cooperative ties among ATL, HSL, MERL, and VGO. The integration of the expertise of each individual lab with those of the others affords more powerful adaptation to overall solutions and best leverages the power of our Peaks of Excellence. Open Community, Volume Graphics, and our Home Networking-related activities are prime examples of how inter-laboratory collaboration produces successful innovation.

HSL, ATL and MERL were founded separately in 1984, 1993, and 1991 (respectively) and each one became a strong individual contributor to MELCO. The three labs were merged together forming ITA during 1995 and 1996 to create an organization with greater capabilities and greater collective understanding of technology trends, market needs, and MELCO business opportunities. One of the first fruits of this combination was VGO, founded in 1998.

Mitsubishi Electric Corporate R&D

ITA is part of MELCO's Global Corporate R&D Network. Corporate R&D, besides supporting MELCO's various business units for new R&D activities, also supports efforts to improve and enhance the range of applications for existing products. The Global Network is made up of six research centers with about 1,800 staff: four labs in Japan, one in Europe, and one in North America (ITA).

The Japanese labs are:

- The Advanced Technology R&D Center (ATC), located near Osaka, works in fundamental areas of materials, devices, and electrical and mechanical engineering.
- The Industrial Electronics & Systems Laboratory (IESL), near ATC, works in such areas as power electronics, factory automation, power plant systems, building controls and automotive systems.
- The Information Technology R&D Center (ITC), near Tokyo works in such areas as information systems, communications, interfaces and opto-electronics.
- The Industrial Design Center (IDC), near ITC, primarily does design and usability work in industrial systems, information and communication systems, and human interfaces.

In Europe MELCO has two laboratories grouped together into Mitsubishi Electric Information Technology Center Europe (ITE):

- The Telecommunications Laboratory (TCL), in Rennes, France does advanced telecommunications R&D aimed at bridging the gap between wireless communication and data networks.
- The Visual Information Laboratory (VIL), in Guilford, U.K. is engaged in research on next-generation digital broadcasting technologies for Digital Audio Broadcasting (DAB) and Digital Video Broadcasting (DVB).

Mitsubishi Electric

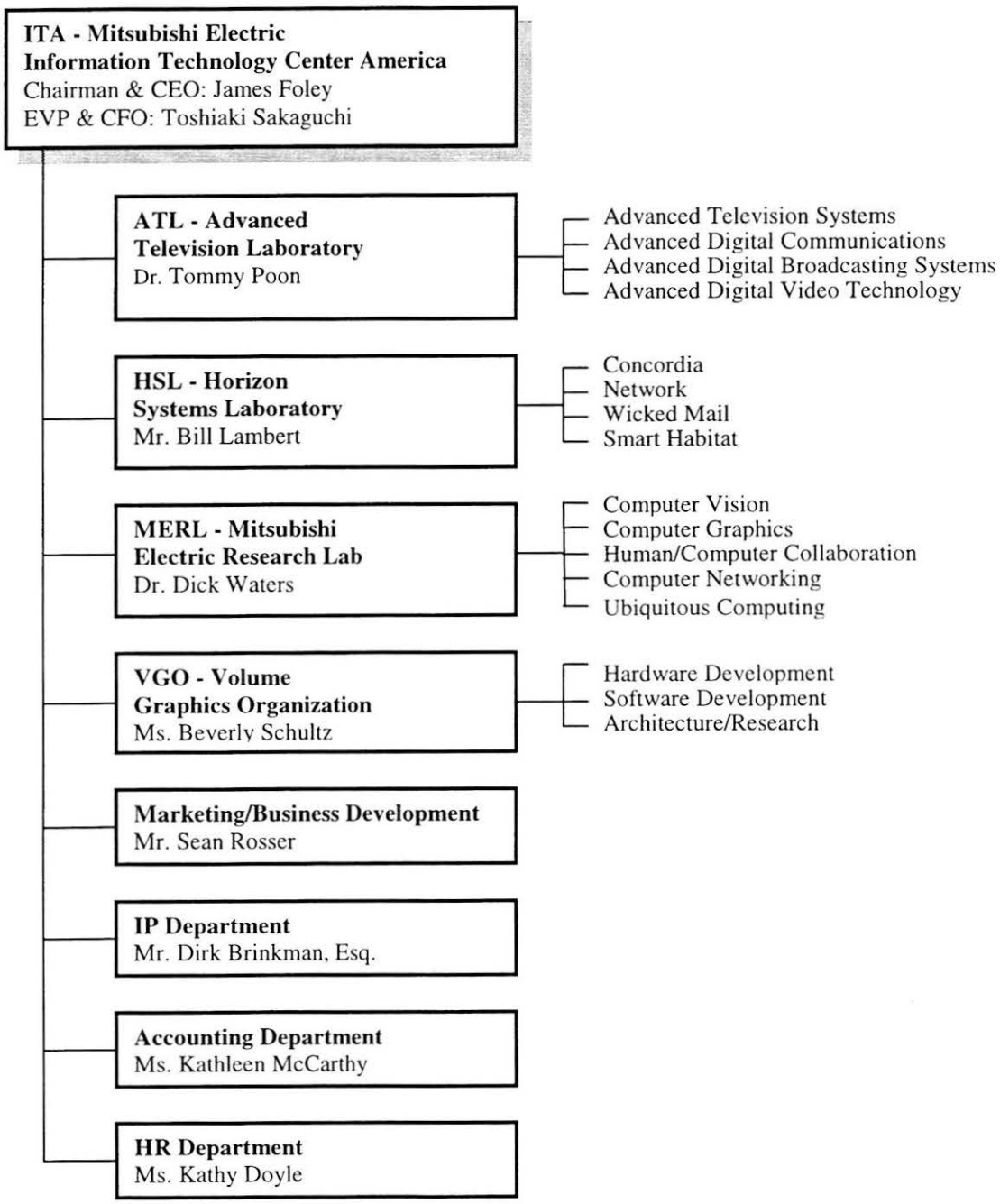
Mitsubishi Electric Corporation (MELCO) is a worldwide corporation with approximately 140,000 employees located in 34 countries. Annual sales for the fiscal year ending in March, 1999, were about $30 billion. MELCO's business is diversified, including semiconductors, communication devices, satellites, computer systems hardware and software, TVs, VCRs, computer monitors, LCD panels, large screen stadium displays, HVAC products, factory automation controllers, robots, power systems and controls, automotive electronics, and elevators and escalators.

(It is worthy of note that there are over 30 companies that use the word "Mitsubishi" in their names and use the red three-diamond logo, including Mitsubishi Motors, Mitsubishi Heavy Industries, Mitsubishi Chemical, and Bank Of Tokyo-Mitsubishi. These companies have shared roots in 19th century Japan, however, these are now complete separate enterprises.)

MELCO business activities in North America include: Mitsubishi Digital Electronics America (projection TVs and home theatre equipment); Mitsubishi Electronics America's Information Technology Group (computer monitors), Mobile Computing Division (the Amity portable computer) and Electronic Devices Group (semiconductors), Mitsubishi Electric Automation (factory automation, robotics), Mitsubishi Electric Automotive America (auto parts), Mitsubishi Wireless Communications (wireless phones and communications infrastructure) and Mitsubishi Electric Power products (electric power products).

Mitsubishi Electric, like many other Japanese companies, has been affected by the lingering malaise of the Japanese economy. Despite this, MELCO's commitment to global and domestic R&D remains strong, with about 5.2% of revenues targeted for this investment in the company's future.

Management Team



James D. Foley

Chairman and Chief Executive Officer of ITA
617-621-7540, foley@meitca.com

Jim Foley's previous positions include: Director of MERL; founding director of the Graphics, Visualization and Usability Center at the Georgia Institute of Technology and Chairman and Professor of Electrical Engineering and Computer Science at George Washington University. He is co-author of three books on computer graphics, an ACM and IEEE Fellow, member of Phi Beta Kappa and Tau Beta Pi, and treasurer of the Computing Research Association. He is a past editor-in-chief of ACM Transactions on Graphics. In 1997 he received SIGGRAPH's bi-annual Stephen Coons Award for his contributions to computer graphics, but the award he most values came from Georgia Tech students: "Most Likely to Make Students Want to Grow Up to be Professors." Foley serves on the board of the Patriot Trails Girl Scout Council. Past research interests include computer graphics, human-computer interaction, and information visualization. He now concentrates on technology management and bringing new technologies to the marketplace.

Toshiaki Sakaguchi

Executive Vice President, Chief Financial Officer and Liaison Officer of ITA
617-621-7513, sakaguchi@meitca.com

As second in command, Toshiaki Sakaguchi is responsible for ITA corporate planning and operation by providing strategic inputs, developing ITA business plans, monitoring budget execution and building strong partnerships with MELCO Japan Business Units, R&D Labs and North American Business Units. He has been a member of the ITA Board of Directors since March, 1996. Prior to his current position, he led R&D projects at MELCO's Industrial Electronics and Systems Lab and Central Research Lab in Japan, being the project manager for the industry-first real-time digital power system simulator as well as the manager of strategic planning and administration.

Tommy C. Poon

Senior Vice President of ITA and Director of ATL
908-302-2814, tpoon@meitca.com

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

William J. Lambert

Executive Vice President of ITA and Director of HSL
781-466-8322, blamb@meitca.com

Bill Lambert's current professional interests relate to how networking and computing are reshaping society. His background is in mainframe systems architecture, and he was president of Horizon Research, Inc for 11 years before it joined ITA in 1995 and was renamed HSL. Under Bill's direction, HSL has made significant contributions to MELCO's product lines in both hardware and software, winning two MELCO President's Awards and two consecutive Corporate R&D General Manager's Awards for excellence. Before joining MELCO, he managed CPU development at IPL Systems and designed mainframes, storage systems, and DSP systems at Cambridge Memories and GTE Sylvania. Outside of work, Bill plays at graphical design and musical composition and occasionally writes humor and fiction.

Richard (Dick) Waters

Senior Vice President of ITA and Research Fellow & Director of MERL
617-621-7508, dick@meitca.com

Dick Waters was a researcher and at MERL from its inception in 1991 until he became director in 1998. His research at MERL centered on multi-user interactive environments for work, learning and play. He led the effort to create the Scalable Platform for Large Interactive Networked Environments (SPLINE) and the Diamond Park demonstration built on top of it. For the 13 years prior to joining MERL, Waters was a member of the MIT AI Laboratory, where he was co-principal investigator of the Programmer's Apprentice project. Waters received his Ph.D. from MIT in 1978.

Beverly J. Schultz

Vice President of Engineering of VGO
617-621-7590, schultz@meitca.com

Beverly Schultz has over 20 years of hardware and software engineering management experience, specializing in multimedia hardware and software and has developed more than 15 successful new technology products. Previously, she was VP of Engineering for Number Nine Visual Technology, moving the company's technology from 2D graphics accelerator boards to leading-edge 3D polygon graphics technology and winning seven technology awards for graphics at COMDEX97. As VP of Software Engineering for Avid Technology, Bev led the production of video editing systems. Several of her people received Academy Awards for Technical Achievement. Beverly was Group Manager of over 250 engineers in several positions at Digital Equipment Corporation, and managed automated factories for Systems Research Laboratories. She holds a B.Sc. in Mathematics from Valparaiso University and a M.Sc. in Computer Science from the University of Dayton.

Sean M. Rosser

Vice President of ITA, Marketing and Business Development
617-621-7502, rosser@meitca.com

Sean Rosser has over 15 years' experience in high-tech industry, mostly in software. Joining General Electric in 1983, he moved through a series of management roles in operations and sales and marketing. After graduating from business school, he joined NovaSoft, a document workflow company, where he became VP Marketing and VP International before leaving in 1993. At the Technology Research Group and Mercer Management Consulting, he led teams in strategic management consulting assignments for clients in technology and telecommunications for over three years. Before joining ITA, he was VP Marketing for New Technologies, Inc., a software start-up involved in high-end graphics software. Sean holds a B.Sc. in Mechanical Engineering from the University of California, Berkeley (1983) and a MBA from Harvard University (1989).

Philanthropy

Through the Mitsubishi Electric America Foundation (MEAF), individual volunteer efforts, and company-sponsored events, ITA has a tradition of commitment to philanthropic activities. ITA's focus, like MEAF, is on education and assisting young Americans with disabilities. In total the ITA laboratories donated and raised about $20,000 in funds for worthy organizations in our communities while volunteering their time and effort to many worthwhile causes.

In 1999 MERL's Philanthropy Committee selected five organizations to receive monetary grants designated for use in programs assisting children with disabilities. Selected programs vary from training camp counselors; contributing to constructing a wheel-chair accessible ramp, and helping subsidize the overhead of Partners for Youth with Disabilities, an organization that matches youth with disabilities with adults with similar physical challenges for a teenage mentoring program. Each grantee applied for and received matching grants from MEAF.

In 1999 HSL's Philanthropic Committee also selected five groups that received grants from ITA and matching grants form MEAF. In addition, the HSL committee members inspired their fellow workers to participate in several hands-on money raising events throughout the year including: "HOPE for Kids" golf tournament, a barn-repair and painting day at LOVELINE (a horse riding program for children with multiple physical disabilities), and a holiday giving event that gathered clothing and needed items for a battered women's shelter.

ATL worked with two charitable organization during 1999: the Somerset Home for Temporarily Displaced Children and the Holy Family Catholic Scholarship Fund.

Over the years, ITA has worked with and supported a variety of organizations including: Outdoor Explorations (enabling children with multiple disabilities to enjoy outdoor activities), EagleEyes-Campus School (providing computer access to children with no motor functions through an innovative eye movement technology), Easter Seals, Boston's WGBH-TV (National Pubic Television), and many other organizations in our employees' neighborhoods.

ATL – Advanced Television Laboratory

The Advanced Television Laboratory (ATL) 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 these challenging problems.

The Digital Networks group is focusing on the design and integration of Home Service Gateways linking external networks to home networks based on IEEE1394, Powerline and Phoneline, and the development of a next-version PSIP product and initial extensions to support data broadcasting. They are also exploring potential applications of MELCO's "Misty" encryption technology to end-to-end solutions. They are active participants and contributors to industry standards for HomeAPI, Echonet, and HomePNA.

ATL has filed more than thirty patents arising from the above activities. Our 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** for MPEG-2, MPEG-4, MPEG-7, XDSL and CDMA; for consumer electronics: ATSC, EIA and IEEE; for home networking: VESA, DAVIC and HAPI.
- **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.

We have developed a close working relationship with three MELCO R&D labs: the Imaging System/Development Center (digital projection TV with 1394 interface), the Information Technology R&D Center (advanced digital broadcasting) and the System LSI Business Center (second-generation video decoder). ATL and MELCO's Semiconductor Business Unit teamed with Lucent Technology on the design of the world's first digital HDTV receiver chip set.

The Advanced Television Laboratory (ATL) was established in 1993 by Mitsubishi Consumer Electronics America (MCEA) and joined Mitsubishi Electric Information Technology Center America (ITA) in 1995.

The laboratory has two locations. The main facility is in northern New Jersey, with easy access to New York City and major airports. We are close to Lucent Technology (Bell Labs), Sarnoff Research Center, IBM Yorktown Heights Research Center, and Princeton and Columbia Universities. ATL has a satellite office in northern California in the heart of Silicon Valley. The current staff consists of nineteen technical professionals, six with Ph.D. degrees and ten with Master's degrees.

Recent Major Publications

P. DaGraca, O. Duardo, S. Hosotani, S. Sugawa and H. Jiang, "A Cost Effective MPEG-2 HDTV System and Video Decoder IC with Integrated Down-Decoder, Graphics and Display Processor", in *Proc. Int'l Conference on Consumer Electronics*, Los Angeles, CA, June 1999, pp. 122-123.

P. De, J. Bao, and T. Poon, "A Calculation-Efficient Algorithm for Decision Feedback Equalizers", in *Proc. Int'l Conference on Consumer Electronics*, Los Angeles, CA, June 1999, pp.64-65.

T. Poon, "Key Enabling Technology for HDTV", in *Proc. IEEE Symposium on Advances in Wired and Wireless Communications*, Princeton, NJ, March 1999.

A. Vetro, H. Sun, and Y. Wang, "MPEG-4 Rate Control for Coding Multiple Video Objects", *IEEE Trans. Circuits Syst. Video Technology*, Vol. 9, No. 1, pp. 186-199, February 1999.

A. Divakaran, H. Ito, H. Sun, and T. Poon, "Scene Change Detection and Feature Extraction for MPEG-4 sequences", in *Proc. SPIE Conference on Storage and Retrevial for Image and Video Databases*, San Jose, CA, January 1999.

A. Vetro, H. Sun, Y.K. Chen, and S.Y. Kung, "True Motion Vectors for Robust Video Transmission", in *Proc. SPIE Conference on Visual Communications and Image Processing (VCIP '99)*, San Jose, CA, January 1999.

Y.K. Chen, A. Vetro, H. Sun, and S.Y. Kung, "Frame-rate Up-conversion Using Transmitted True Motion", in *Second Int'l Workshop on Multimedia Signal Processing (MMSP '98)*, Los Angeles, CA, December 1998.

J. Bao, C. Y. Lu, P. DaGraca, S. Zeng, and T. Poon, "A New Timing Recovery Method for DTV Receivers", *IEEE Trans. on Consumer Electronics,* Vol. 44, No. 4, November 1998.

J.Bao, "A Reconfigurable Adaptive Equalizer for DTV Receivers", in *Intl. Conf. on Signal Processing Appl. and Technology (ICSPAT98),* Toronto, September 1998.

A. Vetro, H. Sun, P. DaGraca, and T. Poon, "Minimum Drift Architectures for Three-layer Scalable DTV Decoding", *IEEE Trans. Consumer Electronics*, Vol. 44, No. 3, August 1998.

K. Tsunashima, K. Muneishi, H. Morikawa, H. Kayashima and V. Sinyansky, "An Integrated DTV Receiver for ATSC Digital Television Standard", *IEEE Trans. Consumer Electronics*, Vol. 44, No. 3, August 1998.

V. Sinyansky, J. Cukier, A. Davidson and T. Poon, "Front-end of a Digital ATV Receiver", *IEEE Trans. Consumer Electronics*, Vol. 44, No. 3, August 1998.

A. Vetro and H. Sun, "Frequency Domain Down-conversion Using an Optimal Motion Compensation Scheme", *Int'l Journal of Imaging Systems & Technology*, Vol. 9, No. 4, August 1998.

A. Vetro and H. Sun, "On the Motion Compensation Within a Down-conversion Decoder", *Journal of Electronic Imaging*, Vol. 7, No. 3, July 1998.

A. Vetro, H. Sun, and Y. Wang, "Joint Shape and Texture Rate Control for MPEG-4 Encoders", in *Proc. Int'l Symposium on Circuits and Systems (ISCAS '98)*, Monterey, CA, June 1998.

A. Vetro, H. Sun, P. DaGraca, and T. Poon, "A Minimum Drift Three-layer Scalable DTV Decoder", in *Proc. Int'l Conf. Consumer Electronics (ICCE '98)*, Los Angeles, CA, June 1998.

K. Tsunashima, K. Muneishi, H. Morikawa, H. Kayashima and V. Sinyansky, "An Integrated DTV Receiver for ATSC Digital Television Standard", in *Proc. Int'l Conf. Consumer Electronics (ICCE '98)*, Los Angeles, CA, June 1998.

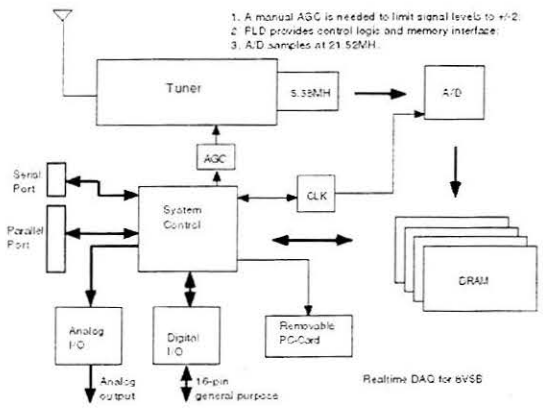
V. Sinyansky, J. Cukier, A. Davidson and T. Poon, "Front-end of a Digital ATV Receiver", in *Proc. Int'l Conf. Consumer Electronics (ICCE '98)*, Los Angeles, CA, June 1998.

J. Bao, C.Y. Lu, P. DaGraca, S. Zeng, and T. Poon, "A New Timing Recovery Method for DTV Receivers", in *Proc. Int'l Conf. Consumer Electronics (ICCE '98)*, Los Angeles, CA, June 1998.

A. Vetro and H. Sun, "Generalized Motion Compensation for Drift Minimization", in *Proc. SPIE Conference on Visual Communications and Image Processing (VCIP '98)*, San Jose, CA, January 1998.

Project Reports - ATL

High Speed Data Acquisition



In system design and implementation of communication receivers, real data testing of the software models and prototypes is a critical step. This is due to the fact that propagation effects are complex and often difficult to model in software, and real data tests early on during system development stage help ensure a successful design. Further more, it is desirable to repeat same test conditions on models which undergo improvements. The high speed data acquisition device described in this project meets such a need. Originally designed for configurable ATV receiver demodulator development, it can be used for collecting data in other high speed transceiver development. A simplified version was prototyped in lab and used for universal modem system development.

ATV: Advanced Television, IF: intermediate frequency, A/D: analog to digital converter, AGC: automatic gain control.

Background and objectives: This project was a part of the universal modem development. The primary objective was to provide a recorded real field data of several tens of seconds for testing the demodulator model and prototype performance. The device should be portable, easy to operate, and interface to a personal computer.

Technical discussion: As shown in the above figure, analog signals from the low IF (5.38MHz) of the RF tuner are digitized by A/D and stored to on board memory. The sampling clock of the A/D is variable, determined by the System Control block. System Control also determines the start and stop of the acquisition process. Typically, over-sampling of factor of 2 to 4 of the base-band data rate maybe needed for digital timing recovery. For this development, the sample clock rate was chosen to be 21.52MHz. A 10-bit A/D (AD876 by Texas Instrument) was used. The analog gain of the input signal is controlled by the AGC block, which is set by System Control. The range of the input signal is limited within 2Vpp. Once the acquisition is completed, the data can be transferred via a PCI or parallel port to a permanent memory device or the host computer's hard disk drive. Various analog and digital interface connectors are available for monitoring and calibration purposes, as well as allowing alternative analog or digital inputs from different sources to be recorded. For the lab prototype, the System Control block is implemented in one Altera FPGA chip, and a SRAM module was used instead of DRAM. For designs which require longer data sequence, DRAM is more economical than SRAM. The DRAM controller was designed in one FPGA. A bus controller chip from Cirrus Logic was chosen for the interface to a PCMCIA portable hard disk. In the universal modem development, data collected in the lab are simply transferred to the host PC via digital I/O (provided by a CY233 interface chip), and later to Unix workstations for analysis.

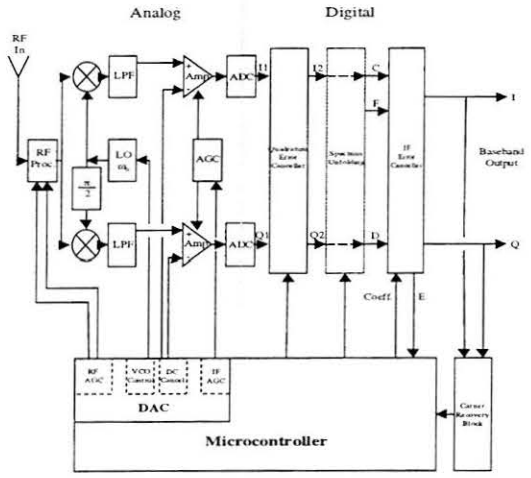
Collaboration: The universal modem project was initially a joint program with Johosoken.

Future directions: With a different RF tuner module, this same device can record data for other transceiver development, such as those used for wireless applications. It is also relatively trivial to increase the number of A/D channels if needed.

Author: Jay Bao

May 24, 1999

Direct-Conversion Tuner for Terrestrial, Cable and Satellite Digital TV



The goal of this project is to develop a new generation of RF tuner for front-end receiver for terrestrial, cable, and satellite for digital TV.

The Direct Conversion Tuner (DCT) will significantly simplify the analog RF sections, eliminate alignment procedures, reduce the size of the RF front-end, and decrease the cost of the entire digital receiver.

In spite that concept of the DCT is not new, only latest advances in digital signal processing (DSP) technology has made it possible to overcome the problems associated with development of a working DCT. New high-speed (DSP) techniques can solve the problems that could not be solved by traditional analog designs.

The concept of the DCT is especially attractive for digital data communications.

Background and objectives: Advantages DCT compare with convention tuner:

- Significant reduction in the number of expensive RF components (tens versus hundred).
- No alignment of the receiver is required.
- Eliminating expensive Surface Acoustic Wave (SAW) filter.
- No image frequencies, no taboo channels.
  High level of integration (implementation with, at most, two chips) easily reprogrammed for different types of modulation (VSB, QAM, BPSK, QPSK, etc.).
- Digital circuits moved closer to the antenna.
- High level of reliability.

Several leading companies such as Motorola, Philips, Maxim, Thomson are working on DCT projects. An IC DCT for satellite TV broadcasting was introduced last time. It is only a question of time when an IC DCT for terrestrial and cable broadcast will appear on the market and completely supplant conventional tuners.

Technical discussion: A proposed block diagram of a DCT is illustrated in the figure above. The incoming RF modulated signal from antenna is bandpass filtered and amplified in the RF Processing block. The output is split into two channels and applied to mixers (I and Q) quadrature channels. These mixers downconvert the signals directly into the baseband frequency. The baseband signals in I and Q channels are digitally processed to cancel DC offset, unbalances and channel unflatness.

Some of the technical challenges for developing a DCT include:

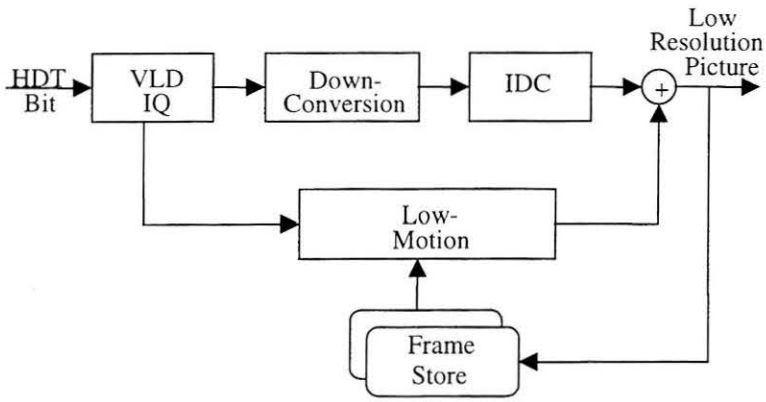
- Amplitude and phase balance requirements between the two quadrature channels (I and Q) must be better than 0.05 dB in amplitude and 0.5° in phase. This is a very tight requirement but can be achieved by correcting the imbalances, which is mostly due to the analog circuitry imperfections, by using a DSP section in the receiver.
- Local oscillator (LO) radiation (leakage) into the antenna input. A matched impedance filter can suppress this leakage with amplifiers on the input and output of the filter.
- DC offset in the signal due to leakage from the LO, analog mixers, and analog amplifiers. This problem can be solved by careful circuit design, digital cancellation, and special modulation of the LO signal.

Collaboration: ATV Lab intends to collaborate with Johosoken for this development as a part of the wireless project.

Future Directions: The idea of DCT can be extended to other types of wideband digital communication receivers. For example, in W-CDMA handsets.

Authors: Victor Sinyansky, Johnas Cukier

HDTV-to-SDTV Down Conversion Decoder



With the current transition to HDTV broadcast, there will be great demand for a low-cost product that is capable of receiving the HDTV bitstream and displaying the signal at a lower resolution. In this project, proprietary algorithms and architectures have been developed, so that our next generation video decoder chip can be competitive. The display devices that we are targeting include existing NTSC television sets and computer monitors.

HDTV: High-Definition Television, SDTV: Standard-Definition Television

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 are applicable to both the TV and PC markets.

Technical discussion: The algorithms for down-conversion and motion compensation have significant impact on the output quality of a low-resolution decoder. In a conventional MPEG video decoder, anchor frames of the video sequence are reconstructed and stored into memory. Using this decoder, a low-resolution signal can be obtained directly from the HDTV signal by decimating outside the decoding loop. Although this method would yield a high quality output, a large amount of memory is required. To overcome this memory requirement, a down-conversion is performed within the decoding loop, hence the motion compensation must be performed from the down-converted anchor frames. The filters used for down-conversion are based on the concept of frequency synthesis, and the filters used to perform the low-resolution motion compensation 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. These algorithms have been conceptually integrated into several realizable architectures and are now being developed as a key feature in the Nozomi II chip set.

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. Since last year, collaboration with Le-Ji-Se has become strong as 5 Japanese engineers have been assigned to the chip set development; Lucent is also involved in this development. Finally, Johoso-ken has provided expert evaluation regarding the quality of the down-decoded video sequences.

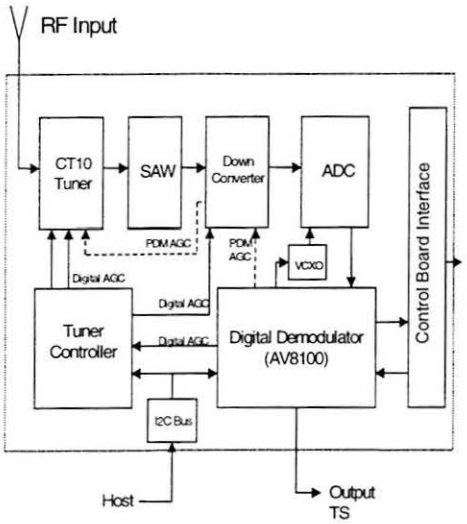
Future Directions: This project has focused on algorithm and architecture development for resolution conversion. We are currently investigating the potential for improvement over the current methods that are being used. Also, we are considering applications that would utilize such conversion techniques, e.g., transcoding and bandwidth utilization within a network as well as decoder implementation on a PC platform..

Authors: Anthony Vetro, Huifang Sun

http://www.meitca.com/ATL/Main.html

May 24, 1999

Front-End of a Digital ATV Receiver



ATV Lab has developed the front-end of a digital ATV receiver. The project included:

- Development of a RF and Digital Demodulator input specification.
- Design of the RF front-end part: a) Tuner, b) Surface Acoustic Wave filter, and c) Downconverter.
- Development of the digital front-end part: a) Analog-to-Digital Converter, and b) On-board microprocessor.
- Development of an AGC algorithm.
- Design of an interface with the digital Demodulator.
- Development and test of evaluation boards.

ATV Lab developed several generations of front-end evaluation boards. These boards were tested in the ATV Lab, in Washington (ATTC), and with real terrestrial broadcast signals in New York (CBS) and in Washington (WHD-TV). The test results were shared with Melco Eijoken and Lucent Technologies. All documentation and sample boards were sent to Eijoken (Kyoto, Japan).

Background and objectives: In North America, the digital HDTV standard was accepted in December, 1996. In the near future, there will be a multi-billion dollar market for Advanced (High Definition) TV receivers.

ATV Lab started this project with the design of a tuner for HDTV. The other part of this project was a development effort on a digital Demodulator chip AV8100 (joint development with Lucent Technology). ATV Lab developed a prototype of the HDTV tuner and tested it together with a Demodulator prototype. As a result of these activities the final concept for a digital ATV front-end was developed. Finally, ATV Lab developed a Test and Evaluation Board and performed extensive lab (ATTC) and field test (Washington DC). ATV Lab has submitted 3 conference reports , and 2 papers describing this project. The achievements of this project were demonstrated at Mitsubishi Open House, COMDEX, NAB, and CES shows.

Technical discussion: A functional block diagram of the front-end of a Digital ATV receiver is illustrated in the figure above. A double conversion tuner, used for cable DTV, according to our specification was modified by the manufacturer in order to increase the gain, the AGC dynamic range, and reduce the noise figure. A total AGC range of 80 dB (UHF) to 85 dB (VHF) was achieved by combining the AGC of the tuner and the downconverter. The AGC data from the demodulator is processed by the on-board microprocessor using an AGC algorithm. For each level of the input signal, the AGC algorithm optimizes the signal-to-noise ratio and minimizes the distortion.

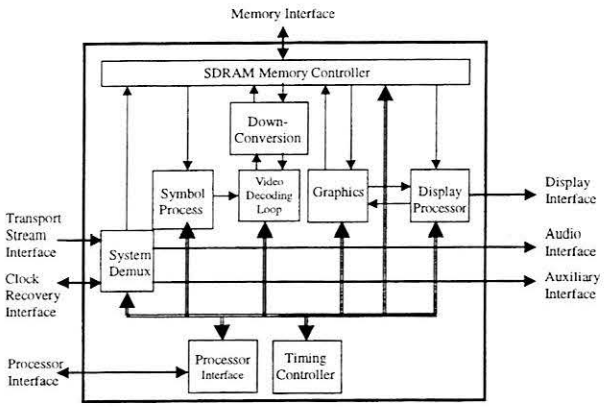
The on-board microprocessor sets up the tuner and demodulator. When a new channel command is received from the host over the I^2C bus, the on-board microprocessor sends a data sequence to the tuner's PLL synthesizers and downloads the new channel's equalizer coefficients to the demodulator. As part of the channel sequence procedure, the demodulator is monitored for "front-end locked", "equalizer convergence", and "valid MPEG2 transport stream output" status. A control board interface was incorporated for the purposes of testing, evaluating, and diagnosing the prototype board. The output of the prototype board is a 188 byte parallel MPEG2 transport stream at 2.69 Mbytes/s for 8-VSB terrestrial broadcast.

Collaboration: Groups inside of Melco include: Eijoken, Johosoken, and CEEC. Groups outside of Melco: Lucent Technologies, Panasonic, Sanyo, Hewlett –Packard, ATTC, WHD-TV, CBS, and MSTV.

Future Directions: Development of a second generation ATV front-end using the direct conversion tuner and reprogrammable type of demodulator chips (Oren, BroadCom, and others). Development of PC-ATV cards using this technology.

Authors: Victor Sinyansky and Johnas Cukier

A 2nd Generation High Definition Television IC



A highly integrated second generation HDTV IC that supports MPEG-2, ATSC, BS4 and other standards is being developed for both HD and SD applications. The IC performs system and video decoding. It incorporates ATL's Down Conversion algorithm for viewing High definition broadcasts on 480I/P displays. For 480I/P displays, internal display processing and 2D graphics is supported. A full resolution output is also provided for HDTV displays. Graphics includes 8-bit color look-up table, translucency, transparency, effects plane, and block draw/copy using graphics windows. Other features include audio and auxiliary data outputs, VCR interface, audio and video clock recovery, and a DNTSC input.

HDTV: High Definition Television, MPEG – Motion Pictures Expert Group, ATSC: Advanced Television System Committee, DNTSC: Digitized Nation Television System Committee, HD: High Definition, SD: Standard Definition

Background and objectives: Over fifty television stations are currently broadcasting digital television signals in the US, however few consumers can afford the high cost ($7000+) of digital HDTV receivers. For the high definition market to flourish, we must design lower cost HDTV receivers. This difficult task requires the development of a highly integrated video decoder solution for High and Standard Definition displays which uses minimum number of memories.

Key Design Features: This IC is able to optimize external memory by using a memory controller that interfaces to 16Mbit SDRAM in two different configurations. For HDTV displays the IC requires six external SDRAM, and for SD or NTSC displays, due to the integrated down conversion technology (developed at ATL) to re-scale input pictures so that a quarter of the external reference pictures is needed, only three SDRAM is required.

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, each with its own data queue.

The video decoding function supports automatic error concealment and timing functions. Timing functions such as 3:2 pull-down for film mode, pause, skip and trick modes can be programmed by an external processor to meet system requirements.

The internal display processor outputs filtered raster video data for 480I/P displays or interfaces to an external display processor of HD displays. It performs 34 different video format conversions. Although the format conversion filters support up to 487 taps with an internal accuracy of 16 bits, all filter have been implemented with only 23k gates using a filter generator (developed by L-Ji-Se).

The IC contains full-featured 2D graphics functions using an 8-bit (256 color) look-up table, translucency, transparency, α-blending, B-blending, Chroma keying, and bit BLT.

In addition, the IC includes audio and auxiliary data channel outputs, a VCR interface, and a DNTSC input.

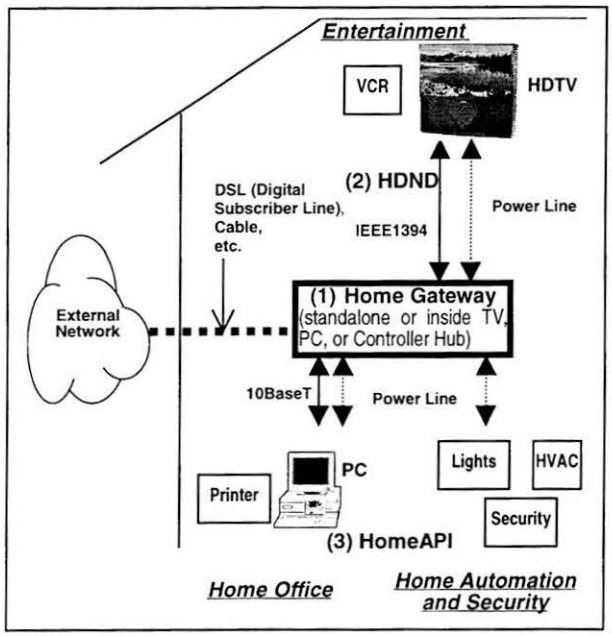
Collaboration: This project is being developed jointly by Mitsubishi Electric - ATL – Advanced Television Laboratory, New Providence, New Jersey, Mitsubishi Electric – System LSI Development Center (L-Ji-Se) Mizuhara, Itami, Hyogo, Japan, and Bell Labs, Lucent Technologies, Murray Hill, NJ.

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

Author: Paul Da Graca

http://www.meitca.com/ATL/Main.html

Home Networking



Digital Home Networks will enable a wide variety of new applications and compelling services in the near future. Their success will heavily depend on providing end users with seamless, easy to use in-home applications, while supporting complex network interactions among in-home devices and outside value-added services. ITA is currently involved in several strategic efforts to enable end-to-end service solutions. **(1) Home Gateway.** Includes the design of a reference platform and prototype for Telephone and Cable Markets. It enables bi-directional flow of Internet Data (e.g. for Web Browsing) and A/V data (e.g. Digital video) between external networks and the home network. **(2) CableLabs® OpenCable HDNI Specification.** ITA co-wrote this specification which defines how HDTV are connected to Cable boxes via a high-speed link (IEEE1394). **(3) Home API Working Group.** MELCO Semiconductor Unit in conjunction with 5 industry leaders is driving this effort. The goal is to specify a common software API to develop client applications able to discover and control home devices connected to Home Networks. ITA supports spec writing.

API: Application Programming Interface, ATM: Asynchronous Transfer Mode, HDNI: Home Digital Network Interface

Background and objectives: Home Networks and Digital CE devices able to *connect* to them can have a great impact in MELCO's CCV businesses. The value of Digital CE devices will depend on the ability to get services from the network. Non-connected devices will become obsolete (e.g. a player without songs is useless). Our main objective is to develop key connectivity technology to support end-to-end services of proven value.

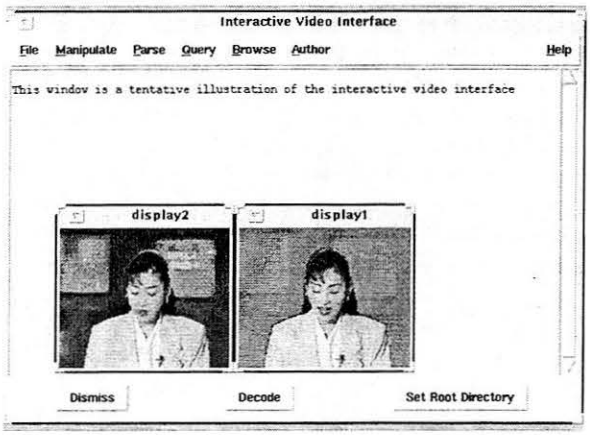
Technical discussion: **(1) Home Gateway** is a key Home Networking component which enables communication between devices connected to the home network and external services. It consists of multiple software and hardware modules supporting a variety of media and protocols. The initial prototype supports several interfaces. In the home network side: IEEE 1394 (for DTVs, PCs, etc.), and Ethernet (for PCs and printers). In the External network side: ATM interface (which can be connected to existing ADSL service provided by US telephone carriers). The gateway distributes AV streams (e.g. MPEG streams for Digital TV) and Internet Data (Internet Protocol packets for Internet applications) to/from devices in the home. The gateway operates reliably 24/7, supports a Web Server, and can be remotely managed. **(2) OpenCable.** ITA co-developed the industry standard for connecting HDTVs to Digital Cable Set Top Boxes via IEEE 1394. In addition to making possible the transmission of High Definition TV programs, the standard specifies: (1) Plug and play capability. HDTVs and Cable Boxes find each other's capabilities when first connected to ensure an easy setup. (2) Graphics transfer feature. The Cable box transmits graphics (e.g. an easy to read TV guide) to be displayed in the HDTV. **(3) HomeAPI.** ITA is co-developing the HomeAPI Specification. The first version of the spec will enable Windows® applications to discover and control home devices. ITA is currently co-developing prototype systems supporting control of devices such as Lights, Air Conditioners, TVs and VCRs via existing power line wires.

Collaboration: **(1) Home Gateway:** Jouhou Souken Multimedia Lab, Cisco. **(2) OpenCable:** Jouhou Souken Multimedia Lab, CableLabs. 1st generation OpenCable HDTV under development by A/V Unit. **(3) HomeAPI:** System LSI Center, Jouhou Souken, and HomeAPI members (Compaq, Honeywell, Intel, Microsoft, Philips).

Future Directions: Commercialize (1), (2), (3). In addition, we started work on **Smart Habitat** concept, which leverages ITA Labs technology on a broader basis: home, office, and road. We intend to combine the experience from the Home Gateway, the company strength in communication and transmission technology (i.e. Wireless, Powerline), and the strategic position as key promoter of industry groups like HomeAPI in a synergistic manner.

Authors: Fernando Matsubara

Interactive Television Project



The Advanced Television Laboratory of Mitsubishi Electric Information Technology Center America Inc. developed the interactive video system to cultivate new MPEG-1, 2 and 4 applications. Object-based manipulation and retrieval have been implemented in software running on the PC. Video analysis and indexing algorithms have been developed as key technologies. Interactive video systems find many applications in digital broadcasting, PC-TV mobile communications etc. The MPEG-7 standard will be integrated to further improve the retrieval performance as well as to enable semantic queries. The left screen shot shows the ability to change the color of an object using MPEG-4.

Background and objectives: MPEG2 has been successfully adopted in digital broadcasting and computer video applications. Now the coding technologies are evolving from MPEG2 to MPEG4 which standardizes algorithms and tools for flexible representation of audio-visual data in an object-oriented manner. The objective of this project is to investigate possible applications of this new coding technology for the future communication environment.

Technical discussion: The object-based compression in MPEG4 enhances the user's interaction with a device or computer. This aspect has been investigated in the project and a demo version of the interactive video system is under way.

The above screen shot shows a window of our demo system. An example of object-based video manipulation is to change the color of an object. In this example since the foreground and the background have been encoded separately, we can change the color of the background easily. Any signal processing such as filtering can be done on an object basis. Interactivity can be at different levels. When we define this term as the capability to allow a sequence of commands and responses, from an application point of view, we can categorize the interactivity into three different levels; program selection, program participation and program creation. Some forms of program selection have already been implemented as in TV channel selection or are going to be available as in the video-on-demand systems.

To achieve a higher level of interactivity more advanced technologies are necessary so that both the commands and responses can be more sophisticated. These include video editing, content-based retrieval and indexing which in turn require scene analysis and feature extraction. We have recently developed a cost effective scene change detection algorithm for MPEG-4/2 coded bitstreams. This will become one of the key elements for development of our interactive video system.

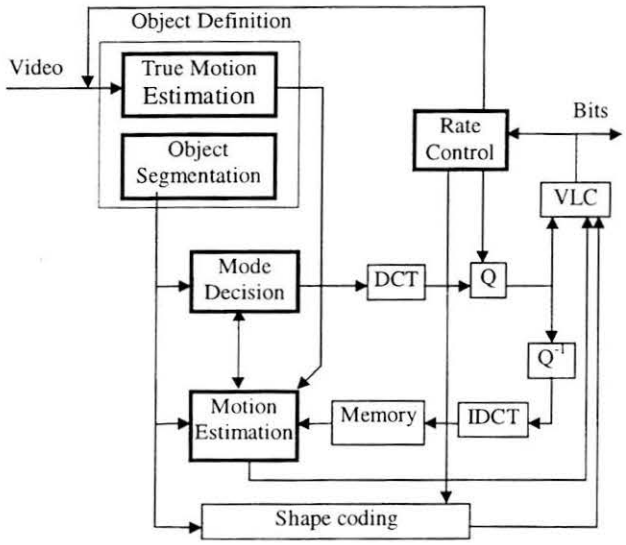
Collaboration: Information Technology R&D Center, Multimedia Information Coding & Transmission Technology Department.

Future Directions: Extend demo system to MPEG7 (Multimedia description interface) as well as enhance and refine current capabilities.

Authors: Ajay Divakaran, Hiroshi Ito, Huifang Sun

May 24, 1999

Object-Based Video Coding Technology



This year, MPEG-4 will become the first international standard to address object-based video coding. In collaboration with Princeton University, ongoing research is being conducted to develop components that are essential to the systems functionality and performance. The blocks outlined in bold indicate current research areas. The work done in many of these areas have sparked new interest within the research community. As these areas gain momentum, it is very important for us to generate intellectual property. This will allow us to have a leading edge in developing products that use this technology. Some possible applications include mobile communications, video telephony, video conferencing, and multimedia applications over the internet.

DCT: Discrete Cosine Transform, VLC: Variable-Length Coder, Q: Quantizer, IDCT: Inverse DCT, Q^{-1}: Inverse Q

Background and objectives: The goal of this project is to develop a set of algorithms for an object-based coding system. Since 1996, the basic areas of research have included motion estimation, object segmentation, rate control and mode decision. Our current emphasis is overall system optimization and software implementation.

Technical discussion: The emerging MPEG-4 standard specifies the syntax and semantics that are needed to decode a compliant object-based bitstream; the techniques used for encoding are fairly open. A rate control scheme to support the object-based functionality of MPEG-4 has been developed. This scheme has been adopted by the MPEG committee and integrated into the software standard. The rate control for multiple video objects is unique in that shape information must also be coded along with texture and motion information for each object. Additionally, a true motion estimation algorithm has been developed. This technique is used to optimize the rate and provide a more accurate estimation of movement within the 2D-image sequence plane. The derived motion field has also shown application for improved frame-rate up-conversion and error concealment in the decoder. To save computation, the motion estimation is considered in a multi-resolution framework. The techniques that are used for object segmentation can be broken down into several steps: i) motion field extraction and principal component analysis, ii) image space ordering and iii) iterative surface optimization. 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.

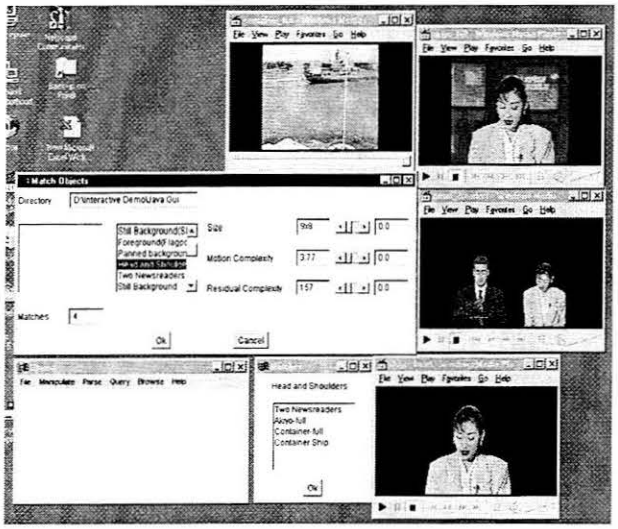
Collaboration: Princeton University has been a significant partner in this work.

Future Directions: The two major directions that are under current consideration are i) improved object-based segmentation algorithm and ii) modeling of the shape information. For the first, we need to test the segmentation of single objects on more test sequences and test the extensibility to multiple video object segmentation. For the modeling problem, it has yet to be determined which shape features best model the rate-distortion characteristics of shape. Preliminary results are pointing towards statistical approaches, such as Markov random fields.

Authors: Anthony Vetro, Huifang Sun

June 3, 1999

MPEG-7 Project



As more and more audio-visual content becomes available in digital form in various places around the world, the ability to locate desired content will become more and more important. Already text based search engines help retrieve textual data from the World Wide Web, but no equivalent identifying information exists for A/V content. The proposed MPEG-7 standard will standardize a multimedia content description interface that will enable efficient searching and browsing of worldwide multimedia content. At the ATL we are developing video indexing techniques to enable rapid browsing and querying . We have proposed a video descriptor to MPEG-7 which is currently part of a core experiment, which if successful will lead to its inclusion in the standard. The screen shot at left illustrates our content based video retrieval system, with the bottom-right image as the user's input

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 on 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. We are therefore focussing on feature extraction from compressed video.

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. We have tested it on the MPEG-7 test set and found that it provides effective matching of shots. We have built a prototype video browsing system using the aforementioned activity features, for both local and remote browsing. Our technology is currently being evaluated in an MPEG-7 core experiment.

Collaboration: This work has been an exclusively ATL effort so far. We are currently investigating possible collaborations with other MELCO labs such as Joho-Soken, Japan and VIL, England.

Future Directions: We plan to add more descriptors to our video browsing framework. We intend to continue our work on feature extraction in the compressed domain. We will pursue inclusion of our current work in the MPEG-7 standard. We will also further enhance the video browsing system.

Authors: Ajay Divakaran and Huifang Sun.

http://www.meitca.com/ATL/Main.html May 24, 1999

Information Security

MISTY chip

In order to support sales and marketing activities of Melco's information security system in US that will be done by MELA/EDG System Marketing, we are proposing our MISTY encryption technology to the academic societies and standardization bodies such as CPTWG and SDMI.

Because of the encryption export restriction, it is difficult for Melco to export their information security products from Japan to US. Therefore, we have developed MISTY encryption chip (CDI 2050) in US in cooperate with Cognitive Designs Inc. and System LSI Development Center. The chip computes hardware encryption and message authentication functions based on MISTY up to 29.3 Megabytes per second.

We are investigating information security activities in US to accelerate Melco's R&D in Japan and other ITA projects in US.

CPTWG: Copyright Protection Technical Working Group, SDMI: Secure Digital Music Initiative

Background and objectives: MISTY was originally developed by Information Technology R&D Center. We are proposing MISTY encryption technology to US academic societies. This project is joint work with Information Security Department, Information Technology R&D Center.

Technical discussion: MISTY is a secret-key cryptographic algorithm developed by Melco. It is designed to operate on a 64-bit data block in Electronic CodeBook (ECB) Mode. It uses a 128-bit key of which all the bits are active (no parity bits inserted). The algorithm is described using an iterative structure. The number of iterations is variable, but must provide a minimum of eight, with additional iterations in multiples of four. MISTY is designed to provide greater security than DES or Triple-Key DES with respect to cryptanalysis. In addition, the 128-bit key size significantly reduces the success of brute-force key search. The structure of MISTY allows fast computational performance in both software and hardware implementations.

We are investigating copyright protection system for Audio/Video using MISTY algorithm and are proposing it to the academic societies such as CPTWG and SDMI. We are also investigating security system in US such as intrusion detection system for Internet, conditional access system for terrestrial broadcasting systm, the electronic commerce system, the digital content distribution, internet payment protocols and the Common Criteria for security product evaluation.

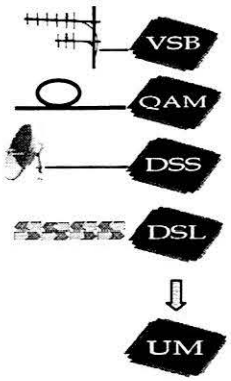
Collaboration: Information Security Dept. Information Technology R&D Center, System LSI Development Center, MELA System Marketing, Cognitive Designs Inc. M-net Dept., Information Technology R&D Center

Future Directions: We have a plan to develop Misty demonstration system using MISTY evaluation board.

Authors: Takashi Okamoto

May 18, 1999

Universal Modem



The rapid advancement of digital multimedia demands more efficient and versatile physical layer solutions. At present, each delivery media (terrestrial, cable, satellite, and telco) requires a unique ASIC chip. In the universal modem project, we aim to develop a single chip, configurable modulator and demodulator (modem) capable of handling digital video, voice, and data signals transmitted via multiple modulation formats. This will significantly increase the flexibility and reduce the hardware costs of modems, which are key elements of the network interface module. Through this project, new and proprietary technology for receiving of ATV signals over the air, cable, and satellite are being developed.

ATV: Advanced Television, ASIC: Application Specific Integrated Circuit

Background and objectives: This project was started in April 1997, jointly with a team at Johosoken led by Dr. Nakayama. It is a strategic advanced development project. The goal for the first stage is to develop new algorithms and architecture for a configurable modem capable of receiving 256QAM, 8VSB, and Q/8-PSK signals. These correspond to the current standards for cable, terrestrial, and satellite ATV broadcast, respectively.

Technical discussion: A modem in a communication device functions as a bridge between the coded information bearing signal and its analog transmission form. In a receiver device, a demodulator must reliably recover the transmitted signal with minimal error. In order to handle multiple modulations efficiently, a new algorithm for the modem is needed. The modem consists of the front-end and back-end portions. A critical part of the front-end portion is the timing recovery circuit, which detects and locks on to the input signal clock based on the information in sampled input signal. A new timing recovery approach, which uses non-synchronized data, has been developed. With this method, all timing information was extracted in digital domain, thus significantly simplifies the timing recovery circuits. In comparison, to realize the same application using a conventional approach with analog components would require multiple voltage controlled crystal oscillators and digital to analog converters, etc. Other major parts of the modem front-end include the carrier recovery and the equalizer. These are designed to meet specifications for different standards, with resource sharing so the logic components can be reused when the modem is configured to work with signals of a different modulation. To achieve maximum hardware efficiency and flexibility, a configurable processor core for the data path was introduced. It was combined with a block-based carrier and gain tracking implemented in software. In addition, a calculation efficient adaptive equalizer that exploits the sparseness of the ATV channel response was developed. Simulation shows that the new equalizer has a convergence speed of an order of magnitude faster than conventional approach, with a smaller final mean square error. To further improve the performance of the receiver under dynamic frequency selective fading conditions, a frequency domain adaptive equalizer was designed.

Collaboration: During the first year of the project, we worked together with Dr. Nakayama's team on specifying the device, complexity analysis, and initial algorithm development. At present, we are in close contact with the team at Johosoken, sharing results and progresses, as most of the output from this project can also been adapted to the multimedia processor chip which are being designed at Johosoken.

Future directions: For the current year, we plan to complete the system design and characterization, we well as partial implementation of key elements of the universal modem capable of receiving VSB, 256QAM and QPSK signals.

Author: Jay Bao

http://www.meitca.com/ATL/Main.html May 24, 1999

HSL - Horizon Systems Laboratory

HSL is MELCO's only laboratory whose mission emphasizes software product development. The future of information technology that HSL incorporates into its vision of computing is one of accelerating digital convergence, ubiquitous computing, and increasing connectivity and user mobility. The importance of information technology is increasing in our daily lives for work, learning, play, and home life. More than just simply technology, information technology is changing the fabric of our social structures in significant ways. HSL is poised to be a prime contributor to the expanding role of Information Technology.

HSL is charged with sensing these fundamental societal changes and 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 laboratory staff with over 300 cumulative years of experience in product development and a leading-edge laboratory infrastructure with an excellent support staff, all located at the heart of the Boston-area high-tech commercial center in Waltham, Massachusetts.

The world requires solutions from information technology to improve the quality of life—increasing the productivity of work, enhancing the spirit of play, and propelling the achievements of learning. HSL is endeavoring to build these solutions either directly or by developing key enabling technologies or tools. We focus on four major solution domains:

- **The Intelligent Enterprise**: the full range of connectivity and integration, from the sales force and E-commerce to the factory floor. HSL provides many software tools in this solution domain: Concordia, Concordia EI, intraAct, Scaleable Mobile Computing, and Wicked Mail.
- **Social Computing**: the real-time connection between people for social applications, such as collaborative working environments, gaming, and networked social spaces. Here, HSL provides Open Community, Schmoozer, intraAct, and Intelligent Traffic Systems Simulation.
- **Enhanced Education**: online learning-related environments such as online education, lecturing, home schooling, and customer support and employee education. Here, HSL provides Concordia, Concordia EI, and Open Community.
- **Smart Habitat**: the home and the integration of home life with trends in ubiquitous computing, mobility and networking. HSL, ATL, and MERL are starting a new effort to provide a platform for digital convergence and connectivity to and within the home. Related technologies are computer vision applications, Concordia, and Open Community.

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

HSL began as Horizon Research, Inc. (HRI) in 1984, a development adjunct of the Information and Communications Systems Division. HRI successfully developed a series of mainframe and office computer systems for the Japanese domestic market. Software development activities began in 1988, and HRI was merged into ITA in 1995, changing its name to HSL. HRI/HSL has received two MELCO President's awards and two Corporate R&D General Manager's Awards, both unprecedented for an overseas development organization.

Recent Major Publications

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

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

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

T. Walsh, N. Paciorek, and D. Wong, "Security and Reliability in Concordia", *in Mobility: Processes, Computers, and Agents*, ACM Press, New York, 1999, pp. 525-534.

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

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

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

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

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

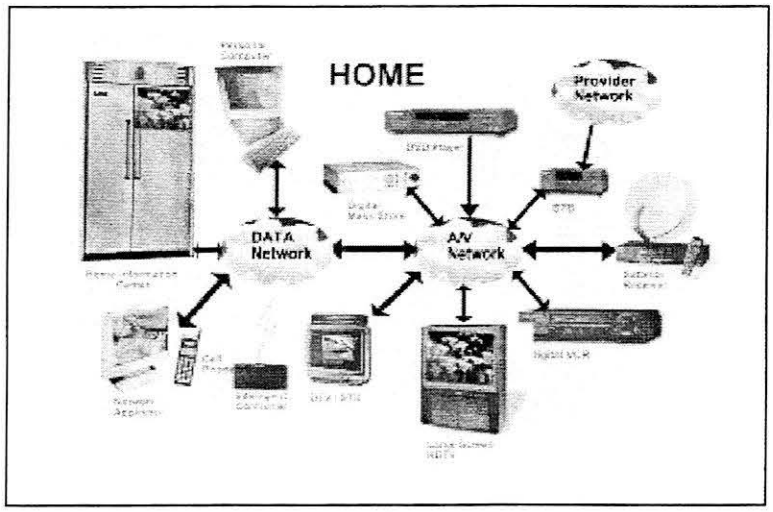
T. Walsh, N. Paciorek, and D. Wong, "Security and Reliability in Concordia", *in Proc. Thirty-First Annual Hawaii International Conference on System Sciences*, Jan. 9-11, 1998, pp. 44-53.

R.C. Waters, D.B. Anderson and D.L. Schwenke, ``Design of the Interactive Sharing Transfer Protocol", in *Proc. IEEE Sixth Workshop on Enabling: Infrastructure for Collaborative Enterprises*, IEEE Computer Society Press, Los Alamitos, CA, 1997.

D. Wong, N. Paciorek, T. Walsh, J. DiCelie, M. Young, and B. Peet, "*Concordia*: An Infrastructure for Collaborating Mobile Agents", Lecture Notes in Computer Science, Vol. 1219, Springer-Verlag, Berlin, 1997, pp. 86-97.

Project Reports – HSL

Smart Habitat



Smart Habitat is a new project effort considering the digital convergence of Home Networking and integration of more and more of Information Technology into daily life.

While we often consider the connectivity between devices, the Smart Habitat effort seeks to present the complexities of a highly connected network structure to the users in a way that they can understand and interface with by having a top level view of the environment as the UI.

Background and objectives: With the incredibly fast growth of the internet and the proliferation of computers in all areas of modern life, technologists are now heralding the imminent arrival of the third wave in computing, which has come to be known as *ubiquitous computing*. In the era of ubiquitous computing, computers will be embedded in devices all around the home and office. These devices will have the ability to communicate and share information with each other freely in much the same way that computers across the world today share information over the World Wide Web. The Smart Habitat project is an advanced research and development effort whose primary goal is to define and develop a unified software infrastructure that will support this interconnection of devices in the networked home environment of the future.

Technical discussion: Current efforts toward ubiquitous computing technology have resulted in the development of various software standards, many of which overlap in their areas of application, and none of which are currently interoperable. The most important of these standards are HAVi ("Home Audio/Visual Interoperability"), HAPI ("Home API"), Jini (from Sun Microsystems) and Echonet. The software infrastructure being developed under the Smart Habitat project will attempt to allow these various and sometimes competing standards to interoperate seamlessly, such that devices that implement one standard will be able to communicate successfully with devices which implement any of the other standards.

Another early goal of the Smart Habitat project is to develop a prototype home environment of the future which will include real devices as well as virtual devices existing under software simulation only. This virtual/real prototype will be used to demonstrate the concepts developed under Smart Habitat, and will in addition serve as a test bed and development platform for applications developed as part of the Smart Habitat effort.

With a technologically advanced product line that extends into many areas of the home and office, Mitsubishi Electric is in a powerful position to participate in defining this new era of ubiquitous computing. Smart Habitat is Mitsubishi Electric ITA's contribution.

Collaboration: ITA ATL, MERL and HSL are all participating in this effort. We hope to have to further cooperation of MELCO labs and business units as this project effort develops, as well as academic partners.

Future Directions: A working prototype of the home and environment of a typical family

Author: Gary Lasker

June 16, 1999

Concordia

Concordia extends on an earlier project to address the needs of the corporate users on the go. Such "road warriors" connect and disconnect frequently with home office servers and thus have strong mobile computing requirements. The Concordia Java-based mobile agent systems framework was developed to address these needs.

Concordia offers a flexible agent mobility model, means for agents to interact, reliability for agent execution and transmission as well as transparent server recovery from failure, complete security features including agent encryption and firewall penetration support, and the most comprehensive administration support among all commercial offerings. Concordia can be tailored to fit the particular hardware needs of the mobile user; only the required components need to be installed. Concordia can be deployed on a spectrum of hardware devices, from smartphones and PDAs to high end backroom servers to fulfill the needs of mobile users for enterprise wide computing.

Concordia™: A Java-based Mobile Agent Systems Framework

Background and objectives: The earlier ***H****ome* ***O****ffice* ***R****oad* project focused on providing solutions for mobile users using other environments as the underlying mobile agent systems platform. However, our experience in this earlier project indicated many shortcomings with existing technology. 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 a flexible model for agent mobility based on its Itinerary construct and was the first commercial product to provide a means for interaction between agents via agent collaboration support. It 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, providing 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's administration support is also the most complete as it provides remote administration and dynamic agent tracking. 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.

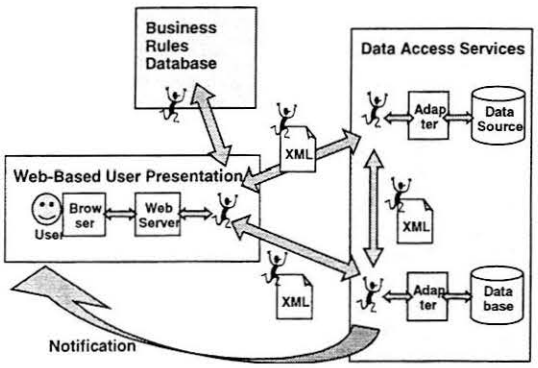
Collaboration: The following organizations within MELCO have used or are currently using Concordia: Sanden (Man-Machine Development Department) for inclusion in the Osaka Waterworks prototyping project as well as new projects for Kansai and Hokkaido Power Electric Companies, Shiden BID for inclusion in the MELBA helpdesk framework, Sentansouken (Cooperative Information Processing Research Team) for inclusion in the Tokyo Electric Power Company's agent negotiation prototyping project. Concordia is also being used in academic research projects at the following universities: Washington University (St. Louis, Missouri), Southern Methodist University (Dallas, Texas), UC Davis Medical School, City College of New York, University of California at Riverside, and University of Cypress.

Future Directions: Current Concordia development has already addressed most of the shortcomings of competing mobile agent frameworks. Future R&D efforts would focus on addressing performance enhancements in basic agent transmission via native TCP/IP transport protocol support (as an alternative to Java RMI), enhancements in application development aids such as better debugging facilities, and development of a "light-weight" version of Concordia that can run on embedded processors such as MELCO's MR32 series controllers.
We also intend to participate in standards activities to ensure wide acceptance of Concordia.

Author: David Wong

Internet Business Initiative Project

Concordia EI Architecture



The Internet Business Initiative Project (also known as Concordia EI) builds upon the Concordia Mobile Agent System, using mobile agents to efficiently and flexibly link together the various enterprise applications and databases systems that make up a company's IT infrastructure.

At its core the Concordia EI 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*. The system provides *data access adapters* which translate information from proprietary formats into XML, and provides a *web-based user presentation layer* allowing users to interact with the system via standard web browsers and internet email.

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

Background and objectives: The Concordia EI 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 EI 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 the Concordia Mobile Agent System and is written entirely in Java. The system operates by converting enterprise information from proprietary back-end formats into industry-standard XML format. Concordia agents 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 adapters* which communicate with existing legacy systems and databases to convert the information contained within these systems into XML format. Such adapters exist for relation database systems such as Oracle, Sybase and Microsoft SQL Server and an adapter to access SAP systems is under development

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 provides technologies that allow users to launch agents, to establish communication sessions with agents and allows agents to render the XML information they carry as HTML pages. The system also supports disconnected access by laptop users via email.

The system's middle tier is composed of a flexible *business rules architecture*. All customer-specific logic that makes up an application is encapsulated as business rules. These business rules are stored in a central repository where they can be easily changed to suit current business policies. Any changes made to business rules are immediately adopted by agents without requiring a system restart.

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.

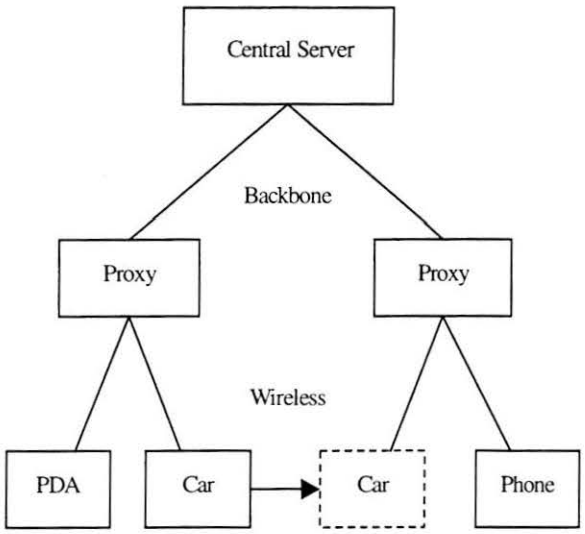
Collaboration: The following organizations within MELCO are currently evaluating the Concordia EI tools for incorporation in their development plans: Sanden (Man-Machine Development Department) for inclusion in the Kansai Electric Power prototyping system and Shiden for inclusion in system integration business opportunities.

Future Directions: Future developments will extend the system to access other enterprise systems such as Peoplesoft or Baan. Development effort may also focus on providing generic application templates that can be quickly customized for a specific customer's needs.

Authors: David Wong, Thomas Walsh
http://www.meitca.com/HSL/Projects/Concordia

June 16, 1999

Disconnected Operation: A Path to the Next Generation of Computing



This MNCRS based project addresses the requirements and solutions for mobile computing. Mobile computing will be one of the main growth streets in the 21st century as information services become mature and reach out to road.

For a long time, wireless connections will continue to be costly, unreliable, and slow. To support mobile users with cost-effective and continuous operation, data that are accessed by the users need to be replicated and updates accumulated on the mobile devices and backbone need to be exchanged and synchronized so that they will be kept consistent at 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 our outstanding contributions to MNCRS, a middleware offering API for applications that need data replication and synchronization support, and a prototype application that demonstrates the functionality and advantages of the middleware

MNCRS: Mobile Network Computer Reference Specification

Background and objectives: In a market research report, the Sales Automation Association estimated that in 1998, 3.5 million salespersons in the United States were equipped with laptop computers and more than 30 million laptop computers are sold each year. Yet in another market report, Garter 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. As can be seen from these reports just about corporate business practice, the demand for disconnected operation will be very common and strong. This forms the ground for the project to develop a middleware that provides general data synchronization support and help enterprise application integration.

Technical discussion: We all just saw the recent craze on Marimba's IPO in the Wall Street. Marimba is a company that develops and sells a product for software distribution on the Internet using one-way replication and synchronization. For more complicated, interactive applications, though, the services for two-way synchronization and concurrent conflict reconciliation along with increased scalability and fault-tolerance must be provided as well. Compared with existing products in the market, our middleware is more superior 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.

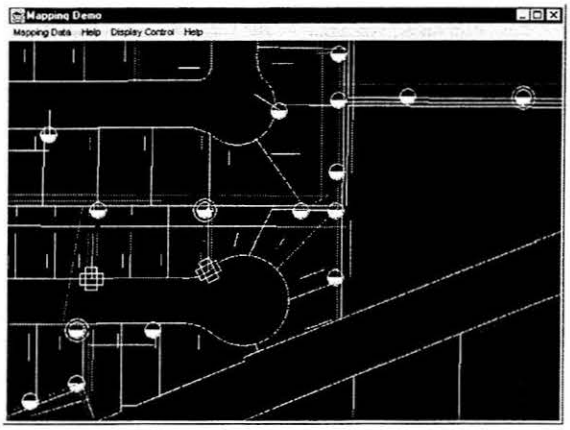
Collaboration: This project has been fully funded by Network Computing Department, Information Systems Laboratory, Mitsubishi Electric Corporation.

Future Directions: The research lasted for 2 years and has finished with an alpha version of the middleware. The beta version of the middleware has been under development and this should be enhanced with industrial performance and quality. After the development of the beta version is finished, the middleware will be able to be used for the NTT DoCoMo project for the company's cellular phone services application development. It can also be the base for any mobile application development inside or outside MELCO.

Authors: Luosheng Peng

June 16, 1999

Electric Power Mapping System



The goal of the EPMS project was to provide a way to develop mapping applications that would run consistently on multiple platforms without requiring customized versions for each platform. The result was a 100% Java implementation of a package that performs high performance display, scrolling, zooming, and selection on vector based maps.

The package can be used in standalone applications or in Web browser applets. It supports a full set of graphic object types as well as user-defined symbols. It is currently being marketed by MELCO as "PreSerV for Web Browser".

AM/FM: Automated Mapping/Facilities Management.

Background and objectives: AM/FM applications for power utilities must be customized for each customer's requirements. These applications typically need to run in a client/server environment on Unix based engineering workstations and on PCs. The goal for PreSerV (presentation server) was to develop a framework for AM/FM applications that can be easily customized, with consistent appearance across platforms, using platform independent source code

Technical discussion: In the initial phase of the project, we evaluated various alternatives for cross platform application development. Java was selected as the basis for development. At this time, Rikijose had a prototype version of a presentation server that was implemented in C/C++ and TCL. This was used as the starting point for development. The prototype was developed around the MD Graphics library, a graphics package that uses MD trees for high-speed search, manipulation, and display of 2D vector graphic information.

Our first prototype added the ability to run Java clients to the existing presentation server, by adding Java network code, a Java class loader, and Java interfaces to the existing C/C++ windowing and graphics functions. In the next phase of development, the MD-Graphics package was recoded in Java, using the Java AWT for display functions.

The package uses double-buffered graphics for improved performance and smooth display scrolling. The object representation of the map is downloaded from a database and kept in local memory. This provides improved performance and reduced network traffic compared to systems where the map image is generated on the server and must be downloaded each time the map is scrolled or zoomed. The local object storage also allows for arbitrary user defined attributes to be associated with each object.

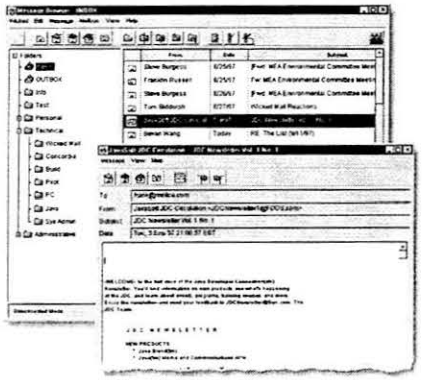
Collaboration: This project was developed in conjunction with Rikijose (Power Information Systems Engineering Center).

Future Directions: A 100% Java version of the software has been completed. Possible enhancements would include updating the software to use the Java 2D Graphics and Swing APIs.

Author: Scott Wilde

June 16, 1999

Wicked Mail – Advanced Internet Messaging



Wicked Mail is a full-featured, next generation email client providing an extremely powerful implementation of the internet standards-based messaging protocol IMAP4 (Internet Messaging Access Protocol). Its feature set targets the mobile computer user, as well as those who need to access their email from multiple locations. It provides advanced synchronization features, allowing the mobile user full capability to read, organize and reply to messages while disconnected from the internet; and then provides fast, complete synchronization of all changes at the next connection. In addition, Wicked Mail Web Edition allows the mobile user full access to all messages and folders from any web browser.

Wicked Mail is written in 100% pure Java, making it suitable for use on a wide variety of computing devices and platforms, including Windows 95/NT, Macintosh, Sun Solaris and Unix, as well as Java-enabled portables and network computers. Wicked Mail is available in both English and Japanese language versions.

Background and objectives: Due to its compelling advantages over the current POP3 standard, the IMAP4 protocol is beginning to take hold in the email world. Market projections estimate the number of IMAP4 users to rise dramatically, surpassing the number of POP3 users around the year 2000. In addition, the number of mobile email users is expected to rise in a similarly dramatic fashion. Wicked Mail squarely targets both of these markets with a unique feature set not seen in any other email product.

Technical discussion: Wicked Mail features several technological advantages over other email clients on the market. Its first important advantage lies in its unparalleled implementation of the IMAP4 internet standard. This includes a very powerful disconnected mode, characterized by an action queuing technology which allows the user to read, delete, respond to, compose, and organize messages, all while disconnected from the internet. Then, at the time of the next connection, Wicked Mail performs a fast and automatic synchronization with the "master" message store on the server. This provides to the mobile and the multiple computer user complete and seamless access to all his or her important messages, anytime and anywhere that Wicked Mail is used.

In addition, Wicked Mail Web Edition provides the mobile user complete access to all messages and folders from any web browser, using the same powerful user interface that is found in the full Wicked Mail application.

Other IMAP-related technological advantages featured in Wicked Mail include a powerful yet easy-to-use server-based search mechanism, advanced inbox filtering capabilities, and exceptional control over download and connect-time parameters. Wicked Mail also has a Japanese language version.

Collaboration: In the past, we have collaborated with the Johosoken NC department, as well as Shiden Mobile NC department. Possible future collaborators include MIND and MSY.

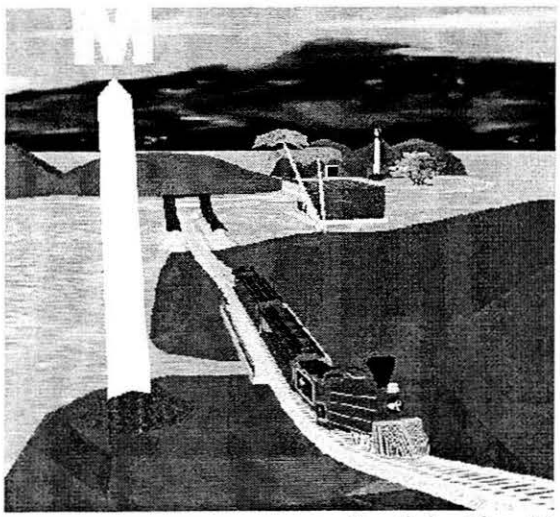
Future Directions: Future plans for Wicked Mail include adding support for the important emerging internet standards LDAP (Lightweight Directory Access Protocol) and ACAP (Application Configuration Access Protocol), further enhancement and refinement of the user interface, further development of the "active mail" concept, and significant augmentation of its Concordia agent-based capabilities.

Author: Gary Lasker

http://www.meitca.com/HSL/projects/wickedMail

June 15, 1999

Open Community (Real Spline)



See Color Plate #1

ISTP: Interactive Sharing Transport Protocol

Real Spline follows the Realtime Collaboration Project as an effort to create a commercial version and robust example applications for Distributed Virtual Environments.

The marketplace is evolving rapidly and Mitsubishi should make a commercial thrust to take advantage of a technology lead in scaleable large systems.

Applications may be used for work, learning or play.

Background and objectives: Real Spline is an effort to make commercial examples of MELCO's Open Community standard, which is based on SPLINE -- A Scaleable Platform for Large Interactive Networked Environments technology. This effort will recognize and increase options to overcome the difficulties in authoring such applications, the limitations and advantages of the platform, recognize partnering opportunities, and generate application examples that demonstrate the power of this new technological field

Technical discussion: Based on the success of making a RTC platform on standard commodity HW, HSL, MERL & Johosouken made a combined effort to improve that platform to the commercial level, and move into the applications and tools space to make applications development an easier task.

ISTP was developed and introduced at the WET/ICE'97 conference at MIT. ISTP is an improvement on scaleabilty and allows distributed realtime collaborative applications to be deployed in a "WWWeb-like" manner: without a central server. In addition, ITA has developed "Halls" application for SigGraph'97, and Schmoozer, an applications authoring and deployment framework with a sophisticated GUI for collaboration and session management.

We improved both audio and visual rendering, as well as improving reliability and error handling across the network. The JAVA API has been completed and full Java dynamic class extension methodology for the framework will be included in the latest work.

Neometron, a venture company, will launch a product based on Open Community later this year, and Johosouken Network Computer Dept. is building the Virtual Medical Center application and industrial applications with MELCO organizations.

Collaboration: Johosouken Network Computer Dept. & Systems Engineering Department and Seisanpon (user of derived applications), Neometron, Nearlife, MSS, MCR.

Future Directions: We intend to release a commercial product and provide a robust platform for MELCO internal applications, as well as seek more outside complimentary partnering opportunities. Open Community will be used as the platform for Schmoozer, our 3D web browsing framework, to be publicly released in 3Q'99.

Authors: Derek Schwenke, Bill Lambert

http://www.opencommunity.com

June 16, 1999

Schmoozer – 3D Web Browser

Schmoozer is a 3D web authoring tool and runtime environment built as a Java framework on top of Open Community.

Users can build virtual worlds from content available on a web server, then flexibly interconnect these virtual worlds to other users, making a seamless and highly scalable virtual world possible.

Users can also collaborate in the authoring environment, as Schmoozer supports multi-user editing.

Background and objectives: Real Spline was an effort to make commercial examples of MELCO's Open Community standard, which is based on SPLINE -- A Scaleable Platform for Large Interactive Networked Enviroments technology. Schmoozer is an example application, a real application, built on top of Open Community.

Technical discussion: Schmoozer is the first application made with Open Community to use the Java API. It provides a framework to support vrml content artists, Java programmers, or application programmers. Schmoozer contains applications authoring functions and a deployment framework with a sophisticated GUI for collaboration and session management.

ISTP -- the INTERACTIVE SHARING TRANSPORT PROTOCOL -- an improvement on scaleabilty and allows distributed realtime collaborative applications to be deployed in a "WWWeb-like" manner: without a central server. Schmoozer uses ISTP to achieve a web server type of service model so that many users can share their virtual worlds and content in the same way that HTTP browsers provide a sharing mechanism. Java dynamic class extension for the framework is supported so that Java developers can build object behaviors and make them available at runtime.

Collaboration: Johosouken Network Computer Dept. & Systems Engineering, Neometron, Nearlife, MSS, MCR, RKC.

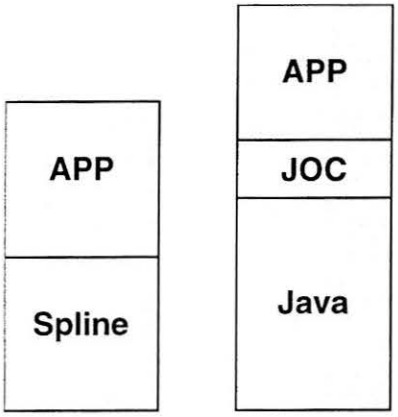
Future Directions: Scmoozer, our 3D web browsing framework, will be publicly released in 3Q'99. We intend to get the feedabck from content makers and VR developers to extend and improve the platform for future releases.

Authors: Derek Schwenke, Bill Lambert

June 16, 1999

http://www.schmoozer.net

Java Open Community (JOC)



Java provides a rich object-oriented model for building distributed systems. In this project we have been studying how to exploit and extend Java's distributed object mechanisms for building scalable, multi-user applications.
JOC programs cooperate via shared Java objects. Objects are actually replicated on different machines, so when one process changes an object all other participating processes see the updated object. Each process is only aware of the objects that are relevant to it and there is no central server, so the system is not limited by the capacity of any machine.
Java Open Community should provide a low-cost platform for the development of large-scale distributed virtual environments, with a low entry barrier for application developers.
Because it is built entirely on Java, JOC will improve as Java does.
By taking advantage of built-in Java mechanisms, JOC offers a powerful programming model in a significantly smaller package than its predecessor (Spline).

APP: application, *JOC*: Java Open Community.

Background and objectives: In the past, we and our collaborators have built distributed, virtual environments in C, using the Spline and Open Community platforms developed at MERL and HSL. Our goal in Java Open Community is to provide similar services to application developers, but to do so with simpler, more powerful mechanisms by taking advantage of Java's object model.

Technical discussion: JOC is an alternative to Spline, so this section compares the two. Some fundamental concepts and the API's for the two systems are rather different. JOC separates control of sharing context from the geometric relationships of objects, replacing the spLocale construct of Spline. Cheap and efficient sharing contexts replace the spBeacon construct of Spline.

JOC advantages:
1. The bulk of Spline is object management logic, eliminated by the use of Java. All of Spline's communication logic is replaced by using mechanisms built into Java. Spline's preprocessor (SPOT) is no longer required.
2. Shared objects may be of almost any arbitrary user-defined class and are not limited to derivations from a provided base-class library. Note, this also permits dynamic classes, not available with Spline.
3. Easier to program with widely used and understood standard Java mechanisms, including object model, threading, class loading. Simpler API.
4. Layered design separates mechanisms from policies to provide more powerful tools to application.

Spline advantages:
1. Extensively optimized for its intended purpose.
2. Already well understood by Spline users.
3. Behavior and performance well understood.

Note that JOC does not yet have 3D logic, audio, or rendering.

Collaboration: We are working with the Open Community and Schmoozer team at HSL, developing an alternative infrastructure to help them achieve their business objectives at reduced cost.

Future Directions: The next major step is to develop or port an application to this new platform.

Authors: David Anderson, Barry Perlman (HSL) June 16, 1999

Intelligent Transportation Systems Development

ITS: Intelligent Transportation Systems, *GIS*: Geographic Information Systems

Background and objectives: The ITS efforts at HSL are focusing on three areas: identifying ITS opportunities in North America, the replication of a driving simulator developed by the Sanken Research, and importing street and road information into Sanken's simulator. The efforts involved with identifying ITS opportunities in North America attempt to provide a context for the development of ITS products by Mitsubishi Electric.

In conjunction with the ITS market evaluation for North America, a driving simulator developed by members of the Sanken Research Staff is being replicated at MERL. The simulator at MERL will serve as a demonstration system for measuring the interest in such devices from North American ITS users and developers.

Technical discussion: Replication of the Sanken ITS Simulator at MERL - The ITS simulator developed at the Sanken Labs must be characterized as a prototype system, and not as a product. Once replicated at MERL, the simulator will serve as a proof of concept system to be evaluated by potential users. If the simulator is to be transitioned into a product, a new development effort will be launched based in part on the concepts and knowledge gained from implementing the original simulator. Completion of the simulator replication at MERL is scheduled for mid July of 1999.

In developing a useful simulator, one of the major challenges is the integration of environmental descriptions into that simulator. In the case of the Sanken simulator, this environmental information consists of street network descriptions, traffic control algorithms for various traffic control mechanisms (stop lights, variable message signs, lane control signs, ...), and vehicular / driver behaviors. Current efforts are focused on importing GIS (Geographic Information Systems) data into the simulator to provide basic terrain and road characteristics. Future efforts will include the capability to create computational models of driving behaviors and traffic control mechanisms.

Collaboration: Dr. Kumazawa - Visiting Researcher at MIT, Sanken ITS Research Group

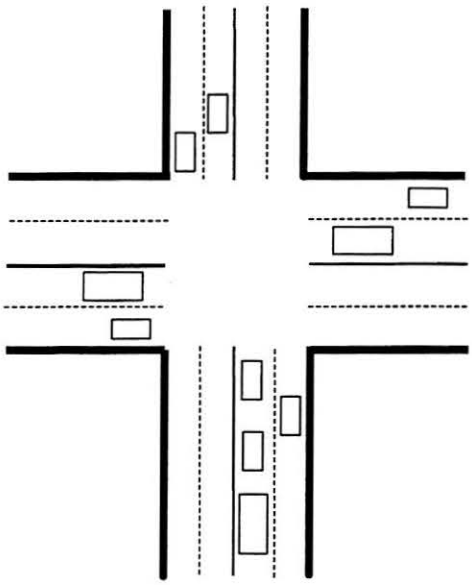
Future Directions: Mitsubishi Electric's collective market knowledge with regard to ITS in North America is rather limited at this point in time. A concerted effort to gain a better market understanding will be actively pursed over the next year. These efforts will include attending trade shows, gathering market data, and establishing appropriate corporate partnerships. In establishing corporate partnerships, Mitsubishi Electric ITA will be pursuing a small group of companies that are the market leaders in applying information technology to ITS in the United States.

The majority of technical effort will be focused on the creation and editing of data to support the ITS simulator. The data creation and editing mechanisms will be designed such that data can be manipulated in a way that is the most natural to the target customer. Over the next year, it is anticipated that the development team will spend much of its time with potential users. The knowledge gained from these domain experts will serve as a solid foundation upon which the data creation and editing tools will be designed and implemented.

Authors: Bill Lambert, Mike Walterman

June 16, 1999

Traffic Control Simulator (TCS)



The Traffic Control Simulator (TCS) is a microscopic traffic flow simulator used to evaluate the performance of dynamic traffic management systems. TCS simulates and provides analyzing tools for various aspects of vehicular urban street and/or expressway traffic.

TCS uses user-defined street network models and provides the capability to display an animation of the simulation as well as several tools to collect and display statistical information about the simulation run. Output is viewed through a Graphical User Interface (GUI) running on X Windows/Motif.

Background and objectives: Economic loss and environmental damage caused by increasing vehicular traffic are serious problems in Japan (and elsewhere in the world). With recent computer and communication systems advances, dynamic traffic management systems based on actual traffic conditions, such as signal control, traffic information services and route guidance, are expected to reduce these problems.

In order to develop new traffic management systems, we must evaluate them under various traffic conditions. Since this is difficult to achieve in the real world due to huge costs and safety concerns, simulation provides an ideal environment for this task.

Technical discussion: Conventional macroscopic simulation models that had been adopted for signal control systems did not deal with traffic on a vehicle-by-vehicle basis. In developing TCS we adopted a microscopic simulation scheme that models each vehicle via a vehicle movement model in order to evaluate signal control systems and other traffic management systems more effectively and accurately. The vehicle model makes "decisions" taking into account vehicle type, road shapes, neighboring vehicles, signal lights, obstacles, etc. The vehicle movement model parameters were derived from observation and analysis of vehicle movement in the real world. TCS modeling can be used to analyze expressway bottlenecks and predict traffic jams, for example. We feel that this microscopic simulation model is a valuable tool to aid urban planners in making traffic systems more efficient.

TCS was written in C++ using X Windows/Motif and runs on Unix platforms.

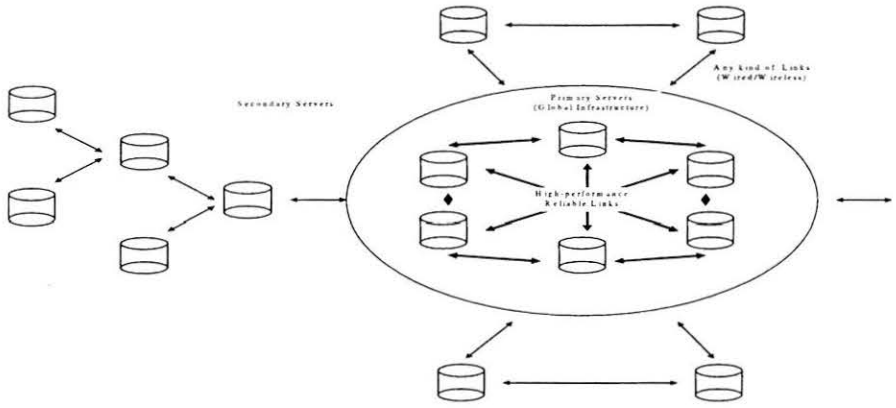
Collaboration: TCS was a joint collaboration between Melco's Industrial Electronics and Systems Laboratory and ITA's Horizon Systems Lab (Horizon Research, Inc at the time).

Future Directions: TCS was followed up by the Next Road Transportation Systems project.

Authors: Bill Lambert

June 16, 1999

PILGRIM: The System for Data Replication and Synchronization over Networks



In the PILGRIM project, we have designed and developed an object-oriented Java-based data replication and synchronization middleware for wide-area desk-top/mobile computing. The middleware complies with MNCRS, and it can support the system topology as shown in the figure above.

Data replication and synchronization has been identified as one of the key technologies for network computing, especially for mobile network computing. To sustain users' continuous and economic activities over networks, the industry must cope with the challenges for system scaleability, disconnected operation, and efficient data communication. The market for such products is just emerging and we have great potential for business success.

MNCRS: Mobile Network Computing Reference Specification.

Background and objectives: As more users began to connect to networks and deploy mobile computers, several issues became recognized. For example, the network connections used by the mobile computers are often less reliable. Therefore, it is necessary to support disconnected operation and adapt to connections with different behavior. The existing systems, however, have assumed the use of primarily reliable connections between locations of replications. The PILGRIM project was then established to invent new technologies which can help cope with the issues.

Technical discussion: The technologies we invented for effectively solving the identified issues include a flexible system architecture, the combination of synchronous/asynchronous data synchronization between peers or client/server, the combination of whole-object/differential synchronization and the automatic switch between the two types of synchronization, a fine-grained, failure-tolerant data synchronization protocol which can help reduce the data being transmitted, and a framework which can control concurrent update and effectively detect/resolve update conflicts. Additional features include persistent storage management, interoperability, and wide applicability. The interoperability is possible primarily because of the middleware's compliance with the MNCRS. Due to all the features described here and elsewhere, our system can be more flexible, reliable and economic than other existing systems in support for working anywhere, anytime.

Collaboration: This project is in collaboration with NC Department, Johosoken, MELCO. We have been also actively participating in the MNCRS working groups.

Future Directions: The alpha version of the middleware is currently under implementation and it will be completed before the end of this calendar year. We also plan to further develop a series of market-oriented applications including collaborative document authoring, software version management and distribution, distributed database management.

Author: Luosheng Peng

June 16, 1999

MERL – Mitsubishi Electric Research Laboratory

MERL 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 ITA 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, researchers at MERL 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 rapidly as possible.

To allowing rapid response to opportunities, MERL 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's current research focuses primarily on five intertwined themes: computer vision, computer graphics, human/computer collaboration, computer networking, and ubiquitous computing.

In computer vision, we focus on real time processing, exploring ways that computer vision can be used as an input modality. An interesting subtheme 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, volume graphics, and graphics-intensive applications such as surgical simulation. Our work on volume rendering led to the creation of VGO and the real time volume rendering chip VGO has just developed.

In human/computer collaboration, our work includes improved computer interfaces with intelligent agents and research related to human learning. Previous work at MERL on multi-user interaction (Spline and Diamond Park) led to the current work at HSL on Open Community and Schmoozer.

In computer networking, we focus on real time interaction, low latency, and the kind of intermittent connectivity that is typical of mobile computing.

The term "Ubiquitous computing" refers to the rapidly growing trend toward embedding computers in special purpose devices rather than using them in a general-purpose form. At MERL, we are particularly interested in devices featuring computer vision and "tangible interfaces"---special computer input devices particularly suited to specific tasks.

Located in Cambridge Massachusetts adjacent to the Massachusetts Institute of Technology, MERL was established in 1991 and joined ITA in 1996. 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. We invite you to learn more about us by visiting our web site "http://www.merl.com".

Technical Staff

Edith K. Ackermann

Ph.D., University of Geneva, 1981

Edith Ackermann is interested in collaborative learning in computer-mediated environments. She focuses on how physical and virtual spaces support people's interactions, and on how people develop senses of identity and community as they meet in virtual worlds. She has pursued these interests in working with technologists, students, teachers, and researchers in milieus concerned with learning and education. Dr. Ackermann was a faculty member at the MIT Media Laboratory, as well as in the psychology departments at the University of Aix-Marseille, France, and the University of Geneva. In Geneva, she collaborated with both Jean Piaget and Berbel Inhelder.

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, he worked on the Mach and Andrew systems projects at CMU, and taught at the Pennsylvania Governor's School for the Sciences. At MERL, 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 researcher in computer vision, interested in the extraction of 3D information from multiple images. He is working on a software library of reusable components to be applied to a range of geometry-based applications. His current work concerns real-time extraction of 3D geometry of faces, to be used for robust tracking of faces in the presence of changing facial expression, occlusions, and changes in illumination.

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 mathematical results include probabilistic theory formation via entropy minimization and signal synthesis via geodesics in density spaces. These techniques are being demonstrated in "digital puppets" -- systems that learn to synthesize realistic and expressive human behavior.

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 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

William T. Freeman is a senior research scientist at MERL, where he studies Bayesian models of perception 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 Gibson

Ph.D., Carnegie Mellon University, 1991

Sarah Frisken Gibson has research interests in the areas of volume visualization, physically based modeling, and surgical simulation. 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 interests include developing a volumetric representation incorporating surfaces, edges, and fine detail to be used as electronic clay in design and CAD/CAM.

Andrew R. Golding

Ph.D., Stanford University, 1991

Andrew Golding is interested in using machine-learning and case-based reasoning methods to develop high-accuracy computer systems, particularly in natural language and speech. More recently he has started working in the domain of intelligent transportation systems. In his Ph.D. research, he developed a novel architecture for combining rule-based and case-based reasoning, work that received the Best Paper award at the 1991 AAAI Conference and the 1993 Gary K. Poock Editorial Award from the American Voice Input/Output Society.

John H. Howard

Ph.D., University of Texas, 1970

John Howard's research deals with storage, concurrency, and distributed computing issues in operating systems. At MERL he has created a file reconciliation system for laptop computers and developed the software for a high-speed ATM network adapter. Previously he directed the Information Technology Center at Carnegie Mellon University and led development of the Andrew File System. He has also studied storage subsystem and hierarchy performance, developed a clustered timesharing system for CDC 6x00 computers, and contributed to the kernel architecture of DEMOS, an early object-oriented operating system. Dr. Howard is Associate Director of MERL.

Ronald L. Johnson

B.A, Louisiana State University, 1990

Ronald Johnson's work focuses on UNIX systems and network administration. He is responsible for the smooth day to day operations of the heterogeneous UNIX environment, and provides guidance on the strategic advancements of the computer and network infrastructure at MERL. Ron has seven years of experience in systems and network administration, and is currently pursuing a Master's degree in Information Technology.

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 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 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.

Neil Lesh

Ph.D., University of Washington, 1998

Neil Lesh's research focuses on developing techniques for making computers better at understanding and collaborating with people. After completing a thesis on scalable and adaptive goal recognition at the University of Washington, he worked as a postdoc with James Allen at the University of Rochester on the TRIPS project. At MERL, he is working on Collagen, an application-independent toolkit for building collaborative interface agents based on discourse theory. Additionally, he is interested in learning and reasoning under uncertainty and has developed new stochastic sampling techniques for monitoring and control of dynamic systems.

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 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.

Neil R. McKenzie

Ph.D., University of Washington, 1997

Neil McKenzie's primary research interest is the interaction between processors and external devices. His dissertation work concerned the design of message-passing interfaces for scalable parallel computers. Also at the University of Washington, he designed and implemented a low-cost functional tester for VLSI chips. In 1981, Dr. McKenzie and International Chess Master Julio Kaplan created the USCF Chess video game for the Mattel Intellivision. In 1990-91, he worked for LaserAccess Corporation to develop IBM PC-based optical disk peripheral systems for IBM mainframes.

Brian Mirtich

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

Brian Mirtich's research interests are in physics-based modeling, computer graphics, robotics, and computational geometry. Much of his work supports rigid body simulation, with applications to engineering and virtual environments. He developed Impulse, a prototype impulse-based dynamic simulator, and V-Clip, a robust collision- detection algorithm for polyhedral objects. He has published widely used code for computing the dynamics of linked rigid-body structures and the inertial properties of polyhedral objects. Dr. Mirtich has also done research in path planning for vehicles with nonholonomic constraints, and synthesis and analysis of robotic grasping strategies.

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, 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.

Egon Pasztor

B.Sc., California Institute of Technology, 1996

Egon Pasztor came to MERL from Caltech, where he studied quantum and classical physics in addition to computer science. His current areas of interest include computer graphics, vision, artificial intelligence, and computer music. He believes that the digital universe has a tremendous potential to help the world and improve quality of life, but we are far from fulfilling this potential because computers today are blind and deaf and unaware of the world as people see it. Egon seeks to change this. In the fall of 1999, Egon will begin graduate school in the Vision and Modeling Group of MIT's Media Lab.

Ron Perry

B.Sc., Bucknell University, 1981

Ron Perry joined MERL 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 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 Paganation 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's research interests include computer graphics, scientific visualization, computer architecture, and VLSI design. He is the chief architect of VolumePro, Mitsubishi Electric's real-time volume rendering system for PC-class computers. 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 Swiss Federal Institute of Technology (ETH) Zurich in 1991.

Erik Piip

Manager, Computer Network Services Group

Erik is the manager of the Computer Network Services Group. The group provides end-user support, and is responsible for the day-to-day operational and technical well being of the lab's computing environment. In addition to dealing with the daily technical and management issues, Erik takes operational responsibility for the networks used within the lab. Erik is also responsible for identifying strategic and tactical lab wide enhancements to the lab's infrastructure, and their ultimate implantation. He has worked for Digital Equipment in multiple roles, ranging from direct service delivery, services support, services planing, server fault management and analysis and server management design. Other interests include using and building large telescopes and instrumentation (CCD cameras) and Amateur Radio (KA1RV, ES5RV and in the past VK2BQP and VK1RV).

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.

Charles Rich

Ph.D., Massachusetts Institute of Technology, 1980

Charles Rich has had a continuing interest in making interaction with computers more collaborative. As co-director of the Programmer's Apprentice project at the MIT Artificial Intelligence Laboratory from 1980 to 1991, he pioneered research on intelligent assistants for software engineers. At MERL, Dr. Rich led the implementation of the laboratory's first multi-user interactive multimedia environment. Most recently, he developed Collagen, an application-independent toolkit based on discourse theory for building collaborative interface agents. Dr. Rich is a Fellow and past Councillor of the American Association for Artificial Intelligence and was Program Co-Chair of the AAAI'98 conference.

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.

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. Her designs for computational media support human cognition and expression while enabling studies of these processes. Research interests include interactive narratives, virtual construction kits embodying principles of math and science, 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 the Department of Architecture and Media Arts and Sciences. She is on the advisory board for the Making Models program of 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.

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) to 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

Chia Shen, Oscar Gonzalez, Krithi Ramamritham, and Ichiro Mizunuma, "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.

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

Neal Lesh, Charles Rich, and Candace L. Sidner, "Using Plan Recognition in Human-Computer Collaboration", in *Proc. of Seventh Int. Conf. on User Modeling*, Banff, Canada, Jun. 1999, pp. 23-32, TR98-23.

Baback Moghaddam, Henning Biermann, and Dimitris Margaritis, "Defining Image Content with Multiple Regions-of-Interest", in *Proceedings of IEEE Workshop on Content-based Access of Image and Video Libraries*, CVPR'99, Jun. 1999, pp. 89-93, TR99-10.

W. S. Yerazunis, D. L. Leigh, W. T. Freeman, and R. S. Bardsley, "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, Mar. 1999, TR99-29.

Baback Moghaddam, Tony Jebara, and Alex Pentland, "Bayesian Modeling of Facial Similarity", in *Advances in Neural Information Processing Systems 11*, M. S. Kearns, S. A. Solla, D. A. Cohn, eds, MIT Press, Jan. 1999, pp. 910-916, TR99-13.

Matthew Brand, "Pattern Discovery via Entropy Minimization", in *Proc .of Uncertainty'99*, Society of Artificial Intelligence and Statistics, Jan. 1999, TR98-21.

William T. Freeman and Egon C. Pasztor, "Learning to Estimate Scenes from Images", *Advances in Neural Information Processing Systems*, Vol. 11, M. S. Kearns, S. A. Solla and D. A. Cohn, eds., MIT Press, 1999, TR99-05.

Hanspeter Pfister, "Architectures for Real-Time Volume Rendering", *Future Generation Computer Systems*, Vol. 15, 1999, pp. 1-9, TR99-18.

Carol Strohecker, Barbara Barros, and Adrienne Slaughter, "Mapping Psychological and Virtual Spaces", in *Online Proceedings of Design Computing on the Net*, sponsored by *The International Journal of Design Computing*, Nov./Dec. 1998. TR98-12.

Matthew Brand and Ken Shan, "Voice-driven Animation", in *Workshop on Perceptual User Interfaces*, Nov. 1998, TR98-20.

Markus A. Schill, Sarah F. Frisken Gibson, H. J. Bender, and R. Manner, "Biomechanical Simulation of the Vitreous Humor in the Eye Using an Enhanced ChainMail Algorithm", *Medical Image Computing and Computer-Assisted Interventions – MICCAI'98*, Lecture Notes in Computer Science 1496, Springer-Verlag, Oct. 1998, pp. 679-687, TR99-27.

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

Sarah F. Frisken Gibson, "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, Oct. 1998, pp. 888-898, TR99-24.

P.A. Beardsley, "A Qualitative Approach to Classifying Head and Eye Pose", in *IEEE Workshop on Applications of Computer Vision*, Oct. 1998, pp. 208-213, TR98-10.

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

Brian Mirtich, "*V-Clip:* Fast and Robust Polyhedral Collision Detection", *ACM Transactions on Graphics*, Vol. 17, No. 3, Jul. 1998, pp. 177-208, TR97-05.

W. T. Freeman, D. Anderson, P. Beardsley, C. Dodge, H. Kage, K. Kyuma, Y. Miyake, M. Roth, K. Tanaka, C. Weissman, and W. Yerazunis, "Computer Vision for Interactive Computer Graphics", *IEEE Computer Graphics and Applications*, May-June, 1998, pp. 42-53, TR99-02.

Andrew R. Golding and Dan Roth, "A Winnow-Based Approach to Context-Sensitive Spelling Correction", *Machine Learning* (Special Issue on Natural Language Learning), Feb. 1998, pp.107-130, TR98-07.

Sarah Frisken Gibson, Christina Fyock, Eric Grimson, Takeo Kanade, Ron Kikinis, Hugh Lauer, Neil McKenzie, Andrew Mor, Shin Nakajima, Hide Ohkami, Randy Osborne, Joseph Samosky, and Akira Sawada, "Simulating Surgery Using Volumetric Object Representations, Real-Time Volume Rendering and Haptic Feedback", *Medical Image Analysis*, Feb. 1998, pp. 121-132, TR97-02.

Baback Moghaddam and Alex Pentland, "Beyond Euclidean Eigenspaces: Bayesian Matching for Visual Recognition", in *Face Recognition: From Theories to Applications*, H. Wechsler, V. Bruce, T. Huang, J. P. Phillips, eds., Springer-Verlag, Berlin, 1998, NATO ASI Series F vol. 163, TR98-08.

William T. Freeman and Paul A. Viola, "Bayesian Model of Surface Perception", *Neural Information Processing Systems*, Vol. 10, 1998, TR98-05.

Brian Mirtich, "Efficient Algorithms for Two-Phase Collision Detection", *Practical Motion Planning in Robotics: Current Approaches and Future Directions*, K. Gupta and A.P. del Pobil, eds., John Wiley & Sons, Chichester, 1998, pp. 203-223, TR97-03.

Charles Rich and Candace L. Sidner, "COLLAGEN: A Collaboration Manager for Software Interface Agents", *User Modeling and User-Adapted Interaction*, Special Issue on Computational Models for Mixed Initiative Interaction, Vol. 8, No. 3/4, 1998, pp. 315-350. TR97-21.

Oscar Gonzalez, Chia Shen, Ichiro Mizunuma, and Morikazu Takegaki, "Implementation and Performance of MidART", in *Proc. IEEEWorkshop on Middleware for Distributed Real-Time System and Service*, Dec. 1997, pp. 2-10, TR97-22.

K. Ryall, J. Marks, and S. Shieber, "An Interactive Constraint-Based System for Drawing Graphs", *Proc. User Interface Software and Technology (UIST 97)*, Oct. 1997, pp. 97-104, TR97-15.

J. Marks, B. Andalman, P.A. Beardsley, W. Freeman, S. Gibson, J. Hodgins, T. Kang, B. Mirtich, H. Pfister, W. Ruml, K. Ryall, J. Seims, and S. Shieber, "Design Galleries: A General Approach to Setting Parameters for Computer Graphics and Animation", *SIGGRAPH 97 Conf. Proc.*, Aug. 1997, pp. 389-400 TR97-14.

Richard C. Waters, David B. Anderson, John W. Barrus, David C. Brogan, Michael A. Casey, Stephan G. McKeown, Tohei Nitta, Ilene B. Sterns, and William S. Yerazunis, "Diamond Park and SPLINE: Social Virtual Reality with 3D Animation, Spoken Interaction, and Runtime Extendability", *Presence: Teleoperators and Virtual Environments*, Vol. 6, No. 4, Aug. 1997, pp. 461-480, TR96-02.

D. H. Brainard and W. T. Freeman, "Bayesian Color Constancy", *Journal of the Optical Society of America*, Vol. 14, No. 7, Jul. 1997, pp. 1393-1411, TR96-38.

Neil R. McKenzie, Kevin Bolding, Carl Ebeling, and Lawrence Snyder, "ChaosLAN: Design and Implementation of a Gigabit LAN Using Chaotic Routing", in *1997 Parallel Computer Routing and Communication Workshop (PCRCW'97)*, Atlanta, GA, July, 1997, pp. 211-223. TR97-08.

Chia Shen and Ichiro Mizunuma, "RT-CRM: Real-Time Channel-based Reflective Memory", in *Proc. IEEE Third Real-Time Technology and Applications Symposium (RTAS'97)*, Montreal, Canada, June 9-11, 1997, pp. 239-251, TR97-01.

Richard C. Waters, David B. Anderson, and Derek L. Schwenke, "Design of the Interactive Sharing Transfer Protocol", *WET ICE '97 -- IEEE Sixth Workshops on Enabling Technologies for Collaborative Enterprises: Infrastructure for Collaborative Enterprises*, IEEE Computer Society Press, Los Alamitos, CA, Jun. 1997, pp. 140-147. TR97-10.

Richard C. Waters and John W. Barrus, "The Rise of Shared Virtual Environments", *IEEE Spectrum*, Vol. 34, No. 3, Mar. 1997, pp. 20-25. WP96-08.

W. T. Freeman and J. B. Tenenbaum, "Learning Bilinear Models for Two-Factor Problems in Vision", in *Proc. IEEE Computer Vision and Pattern Recognition (CVPR '97)*, Puerto Rico, 1997, TR96-37.

Selected Technical Reports

In addition to publishing our results in journals and conference proceedings, MERL produces a series of technical reports. MERL technical reports are typically either documents that are too long for outside publication (e.g., detailed documentation) or preliminary documents describing work not yet published in an outside forum. Some important recent technical reports that have not yet been superseded by external publications are listed below. The full collection of MERL technical reports can be accessed over the web via "http://www.merl.com/reports/".

Harvey Ray, Hanspeter Pfister, Deborah Silver, and Todd A. Cook, "Ray Casting Architectures for Volume Visualization", TR99-17.

John H. Howard, "Reconcile Users' Guide", TR99-14.

Michael E. Leventon and Sarah F. F. Gibson, "Generating Models from Multiple Volumes Using Constrained Elastic SurfaceNets", TR99-09.

Neal Lesh and James Allen, "Simulation-based Inference for Plan Monitoring", TR99-06.

Neal Lesh, Mohammed J. Zaki, and Mitsunori Ogihara, "Mining Features for Sequence Classification", TR98-22.

Neil McKenzie, "Cranium Network Interface: Architecture and Implementation", TR98-14.

Gregory S. Hornby and Brian Mirtich, "Comparing Diffuse and True Coevolution in a Physics-Based World", TR98-11.

Michael E. Leventon and William T. Freeman, "Bayesian Estimation of 3-D Human Motion", TR98-06.

John Howard, Neil McKenzie, Olivier Voumard, and Ross Casley, "Network Latency in DART and Windows NT", TR98-03.

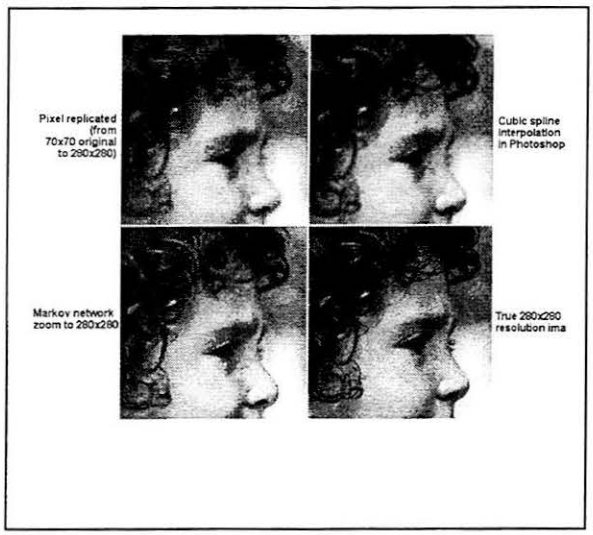
Sarah F. F. Gibson and Brian Mirtich, "A Survey of Deformable Modeling in Computer Graphics", TR97-19.

Carol Strohecker and Barbara Barros, "Make Way for WayMaker", TR97-07.

P.A. Beardsley, "Pose Estimation of the Human Head by Modelling with an Ellipsoid", TR97-06.

Project Reports – MERL

Super-Resolution



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 have developed a Bayesian method to estimate the detail components of a scene, given the low resolution components. The result is shown at left. The input image was the 70x70 image at top left. A standard zoom to 280x280, by cubic spline interpolation, gives the result at the top right. Textures and edges become blurry. The 280x280 output of our new method is at the bottom left. Sharp, linear details in the hair are properly guessed, as is the eyelid. The actual 280x280 image is shown at the bottom right. This new method may have applications to photographic copies, television display, and image enhancement and compression.

Background and objectives: There is much still and moving image content at a low resolution. Much current NTSC programming may be desired to play on future high definition television (HDTV) players. By the time HDTV sets are commonplace, consumers may have come to expect that level of resolution quality. (Just as consumers expect to see color images now, instead of the old black and white ones).

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 describe the patches as vectors in a continuous space, and model the probability densities as gaussian mixtures. (We reduced the dimensionality of the scene and image data within each patch by principal components analysis). We had approximately 20,000 patch samples from our training data, and typically used 9 dimensional representations for both the low-resolution patches (7x7 pixels) and the high resolution patches (3x3 pixels).

Each patch of the low and high resolution images is a node in a Markov network. Given some new image, we seek to infer the corresponding high resolution image components. During inference, we evaluate the prior and conditional distributions of the high resolution data, given the low resolution observation. The high resolution components are a sampling of those high resolution components which correspond to the observed low resolution components at that node. We think of it as a "lineup of suspects". Each node has its own set of suspects. Each scene in a node's lineup has in common the fact that it renders to the low-resolution observation at that node. We evaluate the likelihoods by a set of belief propagation equations. The computation converges in just 3 iterations. The iterations themselves take about 5 seconds each. However, the set-up time prior to beginning the computation takes about 1 hour. We hope to reduce that time with future research.

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

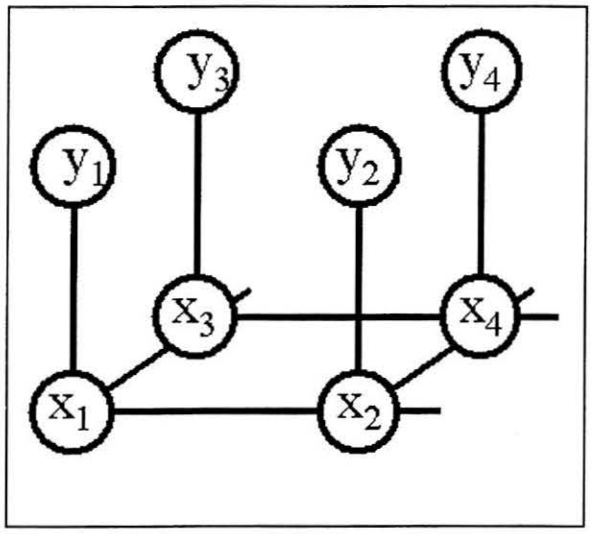
Future Directions: We hope to study the feasibility of this approach for "upconverting" NTSC resolution signals to high resolution signals appropriate for HDTV.

Authors: William T. Freeman and Egon C. Pasztor

http://www.merl.com/reports/TR99-12

June 14, 1999

Correctness of Belief Propagation in Bayesian Networks with Loops



Bayesian networks represent statistical dependencies of variables by a graph. For example, in the figure, the "y" variables may be image values, and the "x" variables may be quantities to estimate by computer vision. Bayesian networks are used in many machine learning applications. Local "belief propagation" rules are guaranteed perform inference correctly in networks without loops. Recently, researchers have found good performance of "loopy belief propagation"--using these same rules on graphs with loops.

We provide theoretical understanding of this good performance. We show that belief propagation converges to the correct posterior means for Gaussian random variables. We show that the convergence points of a related algorithm are at least local maxima 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 a wide variety of applications, including error correcting codes, speech recognition and medical diagnosis. Typically, a probability distribution is assumed over a set of variables and the task is to infer the values of the unobserved variables given the observed ones.

If the graph is singly connected (i.e. there is only one path between any two given nodes) then there exist efficient local message--passing schemes to calculate the posterior probability of an unobserved variable given the observed variables. 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 posterior probabilities (or ``beliefs"). However, as Pearl noted, the same algorithm will not give the correct beliefs for multiply connected networks:

Technical discussion: In this work we analyze belief propagation in graphs of arbitrary topology but focus primarily 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, not just networks with a single loop. The covariance estimates will generally be incorrect but 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.

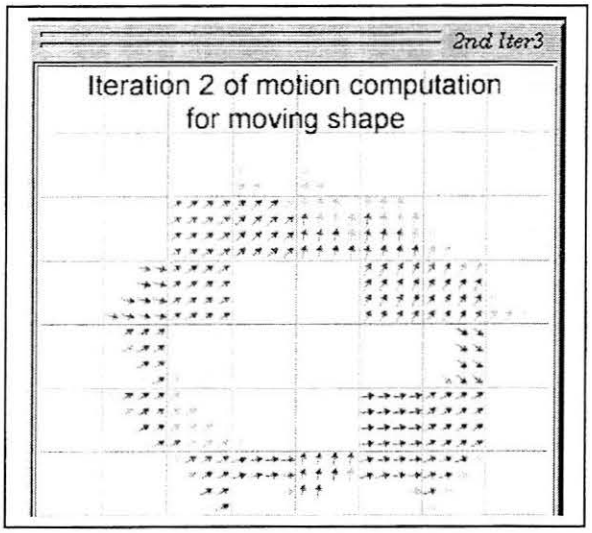
Future Directions: We plan to use these results in applications involving computer vision, such as super-resolution, estimating the shape of an object from its image, and motion analysis.

Authors: William T. Freeman

http://www.merl.com/projects/loopy_learning

June 14, 1999

Learning Low-Level Vision



We have developed a machine learning based method which applies to various problems in low-level vision. We seek the scene interpretation that best explains image data. For example, we may want to estimate the projected velocities (scene) which best explain two consecutive image frames (image).

We use computer graphics to generate synthetic data, and model the statistical relationship between images and scenes in the synthetic world. We then use that model to estimate scene corresponding to a given image.

This yields an efficient method to form low-level scene interpretations, which should apply to a variety of low-level vision problems. We have demonstrated the technique for motion analysis and estimating high resolution images from low-resolution ones..

Background and objectives: We want to get a computer to solve vision problems which are trivial for people: interpret a line drawing; estimate the 3-d shape of an object depicted in a photograph; estimate depth from a stereo pair of image; estimate motion from an image sequence. Of course, algorithms exist for many of these problems, but many are brittle. We seek to exploit the memory capacity of modern computers for solving these problems. We developed a common machine learning framework which applies to all these problems. We hope that learning-based, memory-intensive approach will be more reliable than other algorithms.

There may be many applications of this technology. This research might ultimately lead to a vision chip, which could input image data and output a mid-level scene representation, such as 3-d shape, or reflectances. It might lead to a method to estimate high resolution images from low-resolution ones.

Technical discussion: We ask: can a visual system correctly interpret a visual scene if it models (1) the probability that any local scene patch generated the local image, and (2) the probability that any local scene is the neighbor to any other? The first probabilities allow making scene estimates from local image data, and the second allow these local estimates to propagate.

First, we synthetically generate images and their underlying scene representations, using computer graphics. For example, for the motion estimation problem, our training images were moving, irregularly shaped blobs. Second, we place the image and scene data in a Markov network. We break the images and scenes into localized patches where image patches connect with underlying scene patches; scene patches also connect with neighboring scene patches. The neighbor relationship can be with regard to position, scale, orientation, etc. Third, we propagate probabilities in the Markov network, taking advantage of a "factorization approximation", where we ignore the effects of network loops. This method is fast, and in practise for the problems we have studied, proves to be reliable, as well.

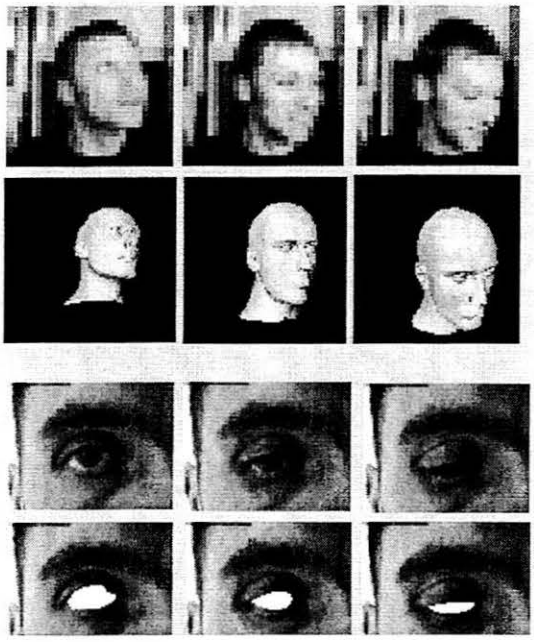
Collaboration: Part of this work, relating to super-resolution, will be in collaboration with the Advanced Television Laboratory in Princeton, New Jersey.

Future Directions: We want to apply the method to more application areas, including: line drawing interpretation, "upconverting" NTSC resolution signals to high resolution signals appropriate for HDTV, motion estimation, and disambiguating shading and reflectance information in images.

Authors: William T. Freeman and Egon C. Pasztor

http://www.merl.com/reports/TR99-12 June 14, 1999

Observing and Classifying the Activity of a Vehicle Driver



The goal of this project is to classify the focus of attention of a vehicle driver, using images from a camera mounted inside the vehicle. One obvious use of this information is for drowsy driver detection (a recent survey suggested that 3% of highway crashes or 100,000 crashes per year in the US are due to tired drivers). But this is only one of the ways to utilize data about the driver's activity. The information can be used in conjunction with readings from external sensors on the vehicle to alert the driver to possible collisions if his or her attention seems to be elsewhere. On a more sophisticated level, it may be possible to learn a driver's characteristic habits, and thereby to anticipate the execution of certain maneuvers. This could be used, for example, to initiate automatic changes to the suspension of a vehicle to facilitate a turn, when observations have indicated that the driver is about to execute that maneuver.

Background and objectives: Today's vehicles have increasing amounts of on-board computer power. Utilizing part of that power to aid the driver - to make driving safer and easier - raises a problem. The driver is already skilled at the task of maneuvering through traffic and crowded urban environments, so a visual or auditory warning to indicate every possible collision would be superfluous and annoying. What is needed rather is the ability to interact with the driver in an intelligent way, offering information only when it is appropriate. This requires observation of the driver, and the ability to understand where his or her attention is being directed. Similar techniques may be of use for other applications such as observation of computer monitor users, enabling hospital patients or those with physical disabilities to use head and eye motions to communicate, and to identify the patterns of attention of consumers in retail stores.

Technical discussion: Since the processor used for driver observation is likely to be relatively low-powered, an approach is adopted which shifts a significant part of the processing to initialization time in preference to run-time. In this initialization stage, a small number of training images of the driver are used in conjunction with a generic 3D model of the head to compute a range of templates, which show the expected appearance of the driver for many head orientations. The run-time scheme matches live images of the driver against the computed templates in order to classify the current head pose. Template matching has the virtue of robustness, coping with images of low resolution (down to 32x32 images for the head pose computation), and is furthermore capable of handling changes in facial expression, localized occlusions, and other localized aberrations such as specularities on spectacles. Once head pose has been determined, processing is focussed on the subject's eyes. Prior models of skin color are used to identify and segment the eye. The shape of the segmented area is used to determine if the subject is looking forwards through the vehicle windscreen, or downwards at the dashboard.

Collaboration: The algorithms in the initial system are suitable for use on the Artificial Retina Module developed by the Advanced Technology R&D Center at IESL.

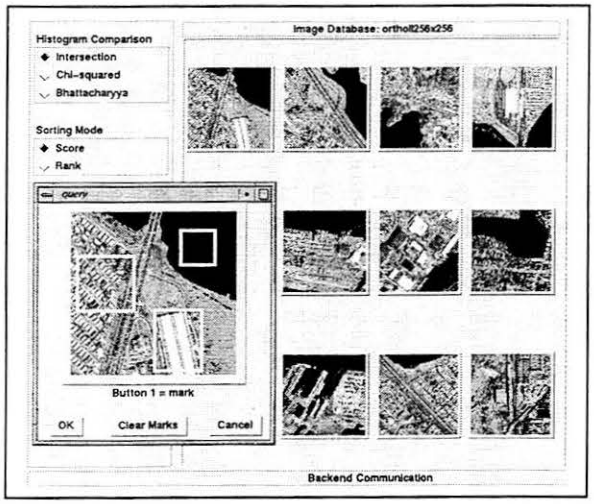
Future Directions: The project above employs a simple generic model for the shape of the human head. New work is focussing on recovering from images the specific 3D shape of each subject's head.

Authors: Paul Beardsley

http://www.merl.com/projects/driverhead

June 7, 1999

Image Retrieval with Multiple Regions-of-Interest



With the proliferation of multimedia, the web and digital imaging, there now exists a high demand for intelligent tools for image management, most importantly indexing, search and retrieval, commonly referred to as QBIC or "query-by-image-content". Existing systems make use of *global* attributes such as overall color distributions which ignore the actual composition of the image in terms of internal structures. The goal of this project has been to develop a new image retrieval system based on the principle that it is the *user* who is most qualified to specify the "content" in an image and not the computer. Therefore the user is asked to provide salient ROIs or "regions-of-interest" and their spatial arrangements in the query image. This technique leads to more acceptable matches returned by the search engine and therefore a more powerful image retrieval tool.

Abbreviations: QBIC: Query-by-Image-Content, ROI: Region-of-Interest

Background and objectives: Most current "query-by-image-content" database indexing and retrieval systems rely on global image characteristics such as color and texture histograms. While these simple descriptors are fast and often do succeed in capturing a vague essence of the user's query, they more often fail due to the lack of higher-level knowledge about what exactly was of interest to the user in the query image - ie. the user-defined content. The goal of this project was to develop and test a new technique in image retrieval using *local* image representations, grouping them into multiple user-specified "regions-of-interest" while preserving their relative spatial relationships in order to build a more powerful search engine for various applications of image database retrieval.

Technical discussion: Image retrieval in general is based on two key components: a set of image features (like color or texture attributes) and a similarity metric (used to compare images). To date most systems use global color histograms to represent the color composition of an image, thus ignoring the spatial layout of color in the query image. Likewise, a single global vector (or histogram) of texture measures (usually computed as the output of a set of linear filters at multiple scales) is used to represent non-color image attributes (such as coarseness, *etc.*) The similarity metric used to compute the degree-of-match between two images is typically a Euclidean norm on the *difference* between two such global color/texture representations. In contrast, our system divides the image into an array of 16-by-16 pixel blocks each of which contains the following feature representations: a joint color histogram in LUV color space and joint 3D histogram consisting of the edge magnitude, Laplacian and dominant edge orientation, computed at two octave scales. These non-parametric densities represent *local* color and texture and due to the additive property of histograms can be easily combined to form bigger regions. When the user specifies a region of interest, its underlying blocks are "pooled" to represent a "meta-block" to be searched for in the database. Multiple regions are likewise searched and the intersection of the best matches determines the final similarity ranking of images in the database. In addition, the user has the option of specifying whether multiple selected regions should maintain their respective spatial arrangement.

Collaboration: Carnegie Mellon University. New York University.

Future Directions: Currently the search for ROIs is computationally intensive and pruning strategies should be implemented in order to avoid searching the entire database for a "meta-block" query. We believe this system will be useful for medical applications where both appearance and spatial factors play a significant diagnostic role.

Authors: Baback Moghaddam

http://www.merl.com/projects/idbr October 11, 1998

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 semiconductor memory, overwriting old video every thirty seconds, and stopping only after an accident triggers the vehicle airbags or other crash sensor.

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

The intended initial market of the CarCam is truck, bus, and taxi fleet owners, and their respective insurance companies. Current fraudulent insurance claims in the US are estimated at 10% of the US $10 billion/year insurance business.

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: A CarCam PEW unit is designed to be mounted in a car, truck, or bus, facing forward, and with a clear viw of the road ahead. The CarCam contains an image sensor such as the Mitsubishi M64283FP image sensor (AR LSI) a CPU with low-power DRAM (such as an M32R), a ROM, a battery, a set of control switches, a small display, a serial or infrared interface, a power-input jack, an accellerometer, and a tamper- and impact-resistant case.

The ROM contains a small control program that repeatedly commands the image sensor to acquire an image, then compresses the image using asymmetrical compression, and stores the image in the low-power DRAM. Successive frames are written until the DRAM is full, then the oldest frames are overwritten. Approximately thirty seconds of video at 5 frames per second are stored; every second or two the PEW indicates on the display that it is active and recording. When a vehicle impact or other accident occurs (as detected by the accellerometer) the CPU continues to store approximately ten seconds of video into DRAM and then either goes into low-power memory-retention mode, or if flash memory is fitted, copies the digital image data into nonvolatile flash memory. The CarCam PEW can retain memory in low-power mode for one week on the internal 9-volt alkaline battery, or indefinitely if using the flash memory.

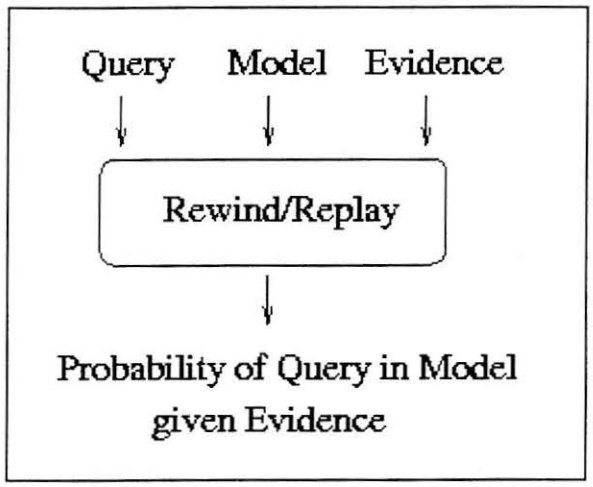
Collaboration: This work was done with the assistance of the department of Neural and Parallel Processing Technology of the Advanced Technology R&D center, with the Network Computing Department, and with the Electronic Devices Group of Mitsubishi Electronics America in Sunnyvale, California.

Future Directions: The current development for CarCam is producing a dozen operable prototypes and demonstrating them to General Motors and Damlier-Chrysler invited technology expositions in the next three months. Our goal is to achieve a high-level business connection with the design and management of these major automobile makers.

Authors: William Yerazunis

June 11, 1999

Simulation-Based Inference



DBNs are a class of Bayesian Networks that model systems that change over time. DBNs have been used to monitor or control dynamic systems such as highway traffic, power plants, and water treatment plants. Because of their size, DBNs are extremely difficult to solve with exact methods or previous approximation methods.

We have developed a new variation of Monte Carlo simulation, called Rewind/Replay, that is effective on large networks. While Monte Carlo simulation is traditionally used only to analyze hypothetical situations, Rewind/Replay can be used to monitor or control dynamic systems given partial observations on those systems. Rewind/Replay can find the most probable cause of the given evidence or project future developments that are likely given the evidence. Our experiments show that our algorithm can solve networks containing over 1000 hidden and 100 evidence nodes in five domains.

DBN: Dynamic Bayesian Networks

Background and objectives: : The original motivation of this work was to use fast, stochastic simulation to perform plan monitoring in large, multi-agent domains. We have generalized our techniques so that they can perform approximate inference on any Bayesian Network. These techniques are especially appropriate for DBNs.

Technical discussion: Stochastic simulation techniques for solving Bayesian Networks require the ability to generate large numbers of random instantiations of the network that are consistent with the given evidence. Previous simulation algorithms perform poorly on DBNs because the evidence tends to be exponentially rare in the number of time steps for which observations exist.

We partition the evidence into small sets and then simulate the network incrementally. We repeatedly simulate the portion of the network that influences the current set of evidence until it is consistent with that evidence. Furthermore, we integrate exact methods into our approach by propagating decisions forward in the network to detect possible conflicts with any of the given evidence that has yet to be matched. Finally, we dynamically re-order the evidence sets in order to increase the effectiveness of the algorithm.

Our algorithm was inspired by a related method in statistics called Sequential Imputation. The only other algorithm we know of that is suitable for large DBNs is the Survival-of-the-Fittest (SOF) algorithm recently developed at Stanford University. Rewind/Replay is significantly simpler and more flexible. Furthermore, some of our techniques can be applied to SOF which improve its performance dramatically. Finally, we have shown Rewind/Replay to outperform (the improved version of) SOF on several example problems.

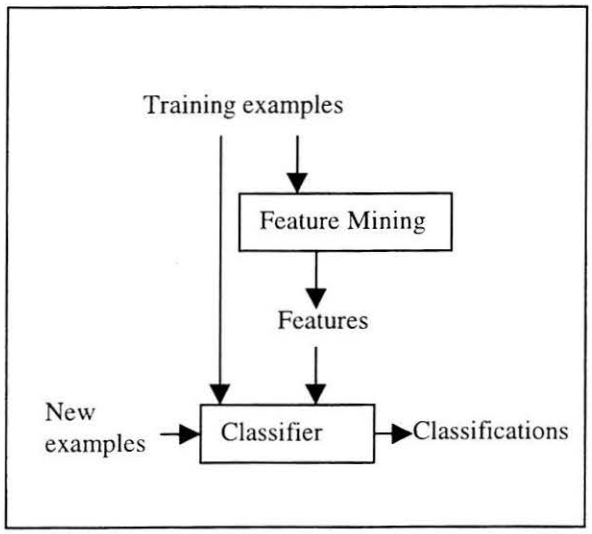
We have developed a wildfire simulation domain. The system's job is to monitor all incoming reports and determine if any fire fighters are in danger of being trapped. We have shown that Rewind/Replay performs better in this domain than previous algorithms.

Collaboration: We are actively collaborating with professors at the University of Rochester and have received sample networks from the University of Saarland, the Ben Gurion University of the Negev, and Aalborg University.

Future Directions: We are planning to apply our algorithm on real problems, perhaps involving electricity distribution or electricity power bidding for deregulated markets.

Authors: Neal Lesh

Mining Features For Sequence Classification



We have applied data mining techniques to the task of feature selection in order to improve the performance of classification algorithms on sequential examples, such as text or DNA sequences.

In the past, it has been difficult to apply classification algorithms to sequential examples because of the vast number of potentially useful features for describing each example. We have adapted data mining algorithms to act as a preprocessor for classification algorithms. Our data mining algorithms search through billions of features and select the ones that are most useful for classification. The rules found by data mining are converted into a set of annotations, which are used to enrich the description of the examples to be classified by the machine learning algorithm.

Background and objectives: This work was originally motivated by the task of monitoring the execution of plans or schedules in order to predict failures before they arise. In this case, there are many features for describing each event and thus an exponential number of features for describing sequences of events. Our approach allows us to apply machine learning algorithms to the task of monitoring temporal processes.

Technical discussion: Some classification algorithms work well when there are thousands of features for describing each example. In some domains, however, the number of potentially useful features is exponential in the size of the examples. Data mining algorithms have been used to search through billions of rules, or patterns, and select the most interesting ones. We have adapted data mining algorithms to act as a preprocessor to construct a set of features to use for classification.

Our algorithm, called FeatureMine, is based on the recently proposed SPADE algorithm for fast discovery of sequential patterns. SPADE is a scalable and disk-based algorithm that can handle millions of example sequences and thousands of items. FeatureMine inherits these properties from SPADE and incorporates several pruning rules that are specific to the task of mining for features.

FeatureMine works by searching through the set of partial descriptions of sequences. The algorithm looks for features that meet the following selection criteria: the features (1) must exceed some minimum frequency in the training examples, (2) must be significantly correlated with at least one of class, and (3) must not be redundant with other selected features.

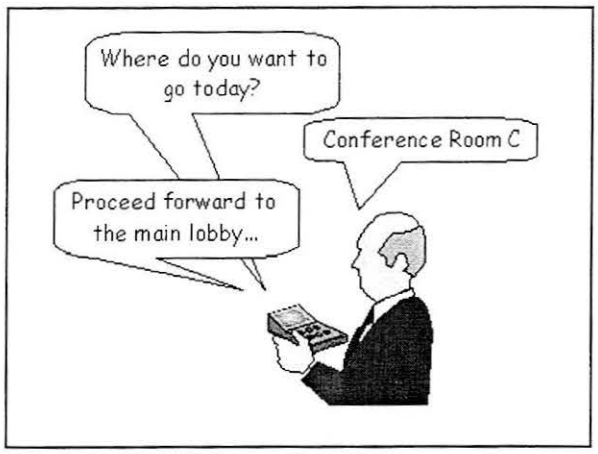
Each pattern produced by FeatureMine is used to create a new boolean features for describing each example. If the pattern holds in the example, then the features value is true, and otherwise is false. We train a standard classification algorithm, such as Winnow or Naïve Bayes, on the enriched examples. Our experiments show that the features produced by FeatureMine improve classification accuracy by 10-50% on several challenging problem sets.

Collaboration: This work is being done in conjunction with the University of Rochester and the Rensselaer Polytechnic Institute.

Future Directions: Future directions include applying FeatureMine recursively on the examples that are incorrectly classified in the first round, and on applying FeatureMine to the task of clustering or unsupervised learning.

Authors: Neal Lesh

Indoor Navigation



This work was motivated by the goal of building a hand-held navigation system that could guide people around urban environments, similar to the car-based navigation systems that are popular today. We have focused on the task of inferring a person's location even in situations in which Global Positioning Systems (GPS) cannot provide this information, such as when the person is indoors or in crowded urban areas where there is no line of sight to the GPS satellites.

As an alternative to installing active badges or beacon systems, we have developed a system which navigates based on naturally-occurring landmarks, such as magnetic fields from steel beams in walls, fixed arrangements of fluorescent lights, and temperature gradients across rooms.

Background and objectives: This work was initiated by discussions about Intelligent Transportation Systems (ITS) at Sanken in 1998. The overall objective is to extend car navigation systems so that they can be removed from the car and carried around by the user.

Technical discussion: In general, we are interested in applying machine-learning techniques to the task of inferring aspects of the user's state given a stream of inputs from sensors worn or carried by a person. We are especially interested in integrating information from the diverse set of cheap, lightweight sensors that are now available, including accelerometers, temperature sensors, and photoresistors.

We have focused on indoor navigation, the task of interactively guiding the user to a desired indoor destination. This task requires, minimally, that the computer be aware of the person's location. For example, the user may be looking for a certain conference room in a convention center, or a train in a large underground train station. Global Positioning Systems (GPS) cannot provide this information indoors or in crowded urban areas. Installing and maintaining beacon architectures involves substantial effort and expense.

Instead, we have explored the other extreme of not modifying the environment at all and using machine-learning techniques to infer a person's location from naturally-occurring signals in the environment. These include characteristic magnetic fields from steel beams in the walls, fixed arrangements of fluorescent lights, and temperature gradients across rooms. Additionally, the user herself provides distinctive acceleration patterns by walking up or down staircases, or riding an escalator or elevator.

We found that the raw sensor signals were unsuitable for use as direct inputs to a machine-learning algorithm. Our navigation algorithm performs very poorly, with almost 50% error, if we use only the raw sensor signals. The reason is that there is too great a distance between the low-level raw signals and the high-level inference we wish to make. To address this, we introduce a "data cooking" module that computes appropriate high-level features from the raw sensor data. These high-level features do not add any new knowledge to the system; they simply reformulate the existing information into a form in which it can be used more effectively by the machine-learning algorithm. Introducing these high-level features was found to improve performance on the indoor-navigation task dramatically. By introducing these high-level features, we are able to reduce the error rate to 2% in our example environment.

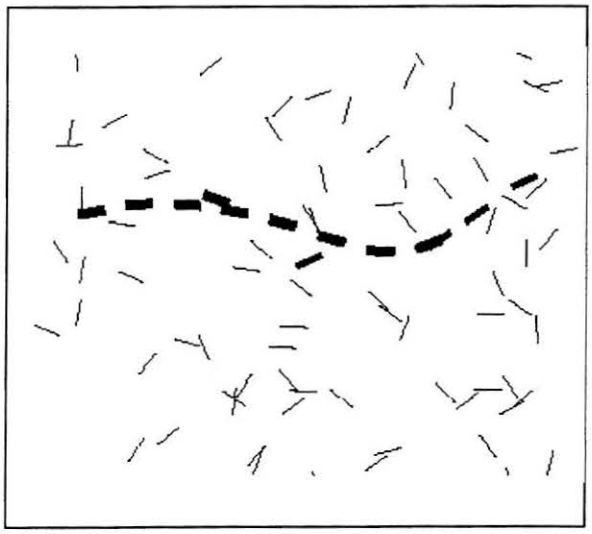
Future Directions: Currently, we are using handcrafted definitions of appropriate high-level features; however, we plan to develop automated procedures for discovering appropriate features through search.

Authors: Andy Golding and Neal Lesh

http://www.merl.com/projects/navi

June 2, 1999

A Factorization Approach to Grouping



Grouping is an important perceptual problem: how do we look at an image and rapidly group the image into clusters of components? This would be a useful task for computers to perform efficiently, having a variety of potential applications in image editing, user interfaces, or image interpretation.

We have developed a method for grouping based on factorizing a particular matrix using SVD. The figure shows algorithm results: for each of the line segments in the image, the saliency of each segment in the discovered foreground group is shown by the line thickness. The algorithm has identified the foreground group, in a time proportional the number of elements grouped.

We explain how this method, using standard mathematical machinery, arises from an axiomatic approach applied to the problem of grouping.

Background and objectives: The figure shows a collection of line elements. It is quite apparent that their mutual positions contain some `global' information. While it is somewhat unclear how to define this global information, it is natural to consider a local property: the `pairwise similarity' of these points: two points that are close by are `similar' and two points that are far apart are `different'. This is a common situation in vision: we extract tokens from an image using some early visual process, and the notion of `closeness' between pairs of tokens is natural and well defined. The tokens may be anything: from pixels to points, to edgels, to textured patches. While it is unclear how to extract, and even how to define, the global high-level properties of the scene, it is easy and natural to define the pairwise affinity of any two tokens. This idea comes to us from the work by Shi and Malik on grouping using normalized-cuts. We explored a simpler approach than that of Shi and Malik: rather than formulating a grouping problem explicitly, we notice that a useful global property of the scene, the foreground set, may be both `discovered' and estimated starting from the notion of pairwise closeness, or pairwise affinity, of individual elements. We have developed a simple algorithm which factorizes the matrix of pairwise element afinities, and it compares favorably with the algorithm of Shi and Malik.

Technical discussion: The foreground group in a scene may be `discovered' and computed as a factorized approximation to the pairwise affinity of the elements in the scene. A pointwise approximation of the pairwise affinity information may in fact be interpreted as a `saliency' index, and the foreground of the scene may be obtained by thresholding it. An algorithm called `affinity factorization' is thus obtained which may be used for grouping.

We have demonstrated the affinity factorization algorithm on displays composed of points, of lines and of brightness values. The affinity factorization algorithm is shown to be computationally efficient (O(n) floating-point operations for a scene composed of n elements) and to perform well on displays where the background is unstructured.

Collaboration: This is joint work with Prof. Pietro Perona, of the California Institute of Technology, Pasadena, CA, USA.

Future Directions: No future work is planned in this area.

Authors: William T. Freeman

http://www.merl.com/reports/TR99-03

May 18, 1999

Learning Concise and Minimally Uncertain Models

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

Entropic estimation is a new mathematical framework that makes it possible to learn the structure and parameters of a probabilistic model simultaneously. 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, on average, faster, more predictive, more general, and much more interpretable than models obtained from conventional learning methods. In many cases, entropic estimation induces a model that is 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, with some parameter twiddling, will fit the 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:

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

$P(\text{data given model})$, called the likelihood, is a known function $f(\boldsymbol{X}, \theta)$, where $\boldsymbol{X}$ is the training data and θ is a vector of parameters. In conventional learning, only the likelihood term is known, so we maximize $f(\boldsymbol{X}, \theta)$. Entropy minimization also gives us the prior,

$$P(\text{model}) = \exp(-H(\text{model}))$$

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 $f(\boldsymbol{X}, \theta)$. 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.

Future Directions: Exact entropic estimators have been derived for most common probability distributions, and many popular models built thereof, including mixture models and hidden Markov models. An approximate estimator has also been derived for generalized recurrent neural networks. We are also looking at learning the structure of Bayes nets. More theoretically, we have established connections between entropic estimation and minimization of expected Kolmogorov complexity.

Authors: Matthew Brand

Voice Puppetry

The voice puppet allows you to 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, automatically incorporates 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 foreign film dubbing.

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), and interpolation between these key poses produces an animation. It is widely understood that phonemes and visemes are really only suitable for lip-syncing, since one loses 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 tracking facial features via computer vision, then modeling their motions 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. This model captures somebody's facial style. 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. We have a fast linear-time closed-form solution for this trajectory.

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 photorealism. 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!).

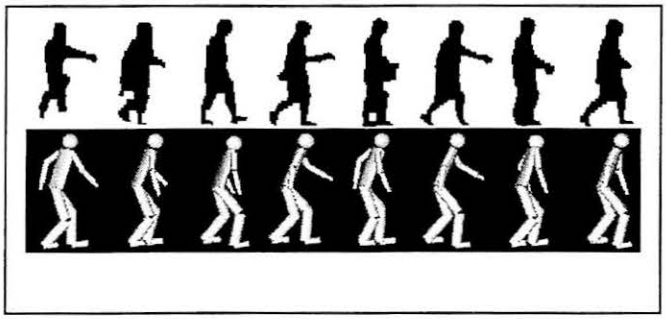
Future Directions: A working prototype was built in 1998. We are fielding licensing inquiries from interested parties.

Authors: Matthew Brand

http://www.merl.com/projects/puppets

May 18, 1999

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.

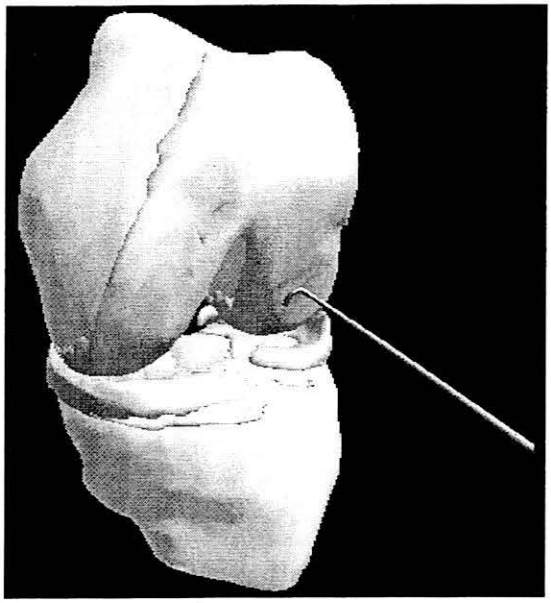
Future Directions: A limited but working prototype was built in the spring of 1999. Current work centers on making the system more robust and synthesizing motion in a variety of styles.

Authors: Matthew Brand

http://www.merl.com/projects/puppets

May 18, 1999

Knee Arthroscopy Simulation



We have developed a knee arthroscopy simulation system that incorporates visualization, haptics, and computer modeling as a research testbed for new technologies in surgical simulation. The most recent prototype uses models generated from hand-segmented high resolution 3D Magnetic Resonance Image data. Volumetric models are used for computer modeling and for haptics while decimated surface models are used for interactive rendering using polygon rendering hardware on an SGI Onyx2 with Infinite Reality graphics.

The most recent prototype allows users to probe the surface of bones in the knee joint with force feedback using a haptic interface device. The device includes a commercial device with 3 degrees of freedom and a custom addition that provides an additional two degrees of freedom, allowing the user to feel forces and torque along the entire shaft of the surgical probe.

The system was demonstrated at the IEEE Visualization conference in October, 1998.

See Color Plate #2

Background and objectives: Surgical simulation has the potential to enhance surgical training, to reduce patient risk by enabling pre-surgical planning on patient-specific models, and to provide intra-operative guidance during difficult surgical procedures. However, many technical challenges remain before surgical simulators can achieve this potential. The goal of this collaborative project is to develop a testbed for new algorithms and technologies in surgical simulation. Knee arthroscopy was chosen as a basis for the system for a number of reasons, including the prevalence of knee surgeries in the United States, the restricted volume of the knee joint, and the importance, but limited volume, of soft tissue in the knee. Our goals include developing methods for high quality visualization, realistic haptic interactions, and dynamic modeling of deformation, cutting, and repair of soft tissues.

Technical discussion: Models for the knee system were generated from hand-segmented Magnetic Resonance Image data. Four data sets have been acquired and segmented from the same subject. They include two sagittally sectioned images, an axially sectioned image, and an image volume acquired for a bent knee. Smooth surface models were generated from the segmented data using SurfaceNets, a technique developed at MERL for this project. These smooth surface models were then used to generate decimated triangle models for rendering, and distance-based volume models for detecting collisions between object models and for calculating probing forces for haptic feedback.

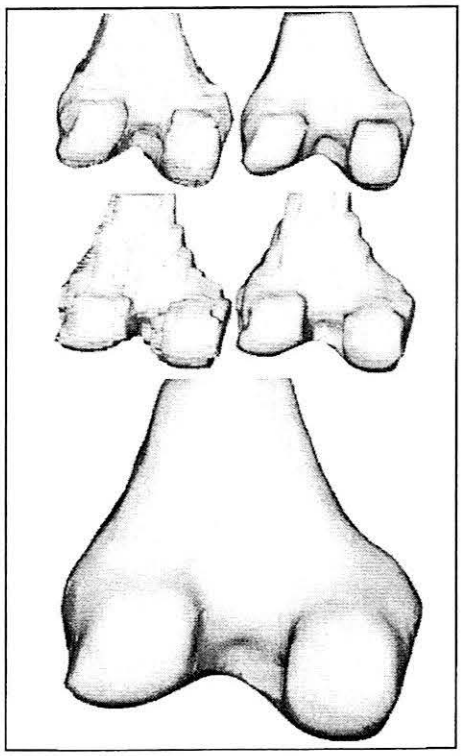
To provide a more realistic interaction for the user, we have built an interface device which consists of a plastic knee model to help orient the user, an arthroscope with a telescoping mount whose position and orientation is tracked and used to control the camera of the graphics rendering system, and a force feedback device that provides haptic feedback. The force feedback device consists of a commercial Phantom from Sensible Devices and a custom device that adds two degrees of freedom to the Phantom's 3 degrees of force feedback. The combined device allows the user to feel forces and torque all along the shaft of the surgical probe rather than just at the tip of the probe.

Collaboration: Massachusetts Institute of Technology's Artificial Intelligence Laboratory, Carnegie Mellon University's Robotics Institute, Brigham and Women's Hospital's Surgical Planning Laboratory

Future Directions: The next step in this project will be the incorporation of soft tissue models into the arthroscopy simulator. Models of articular cartilage and menisci that can be probed, deformed and cut are under development. In addition, we are currently investigating methods for non-invasive acquisition of the mechanical properties of cartilage so that patient-specific models of cartilage could be incorporated into the simulation system.

Authors: Sarah Frisken Gibson

Dual SurfaceNets: Generating Models From Orthogonal Data Volumes



Three-dimensional anatomy models are used in medical diagnosis, treatment, surgical guidance, and surgical simulation. Limited resolution of medical scans can cause artifacts in the models due to low sampling rates between the image slices of a scan. The top four images on the left show these artifacts on a bone model created from segmented MR images with 1mm resolution within the image planes and 5 mm between planes. In the top image pair, the MR images were acquired along horizontal planes. In the bottom pair, the images were acquired along vertical planes. The two models on the left were generated using Marching Cubes, a standard method for generating 3D models from image data. The two models on the right were generated using SurfaceNets. While the SurfaceNet models show a marked improvement over Marching Cubes, artifacts due to the large spacing between image planes are still significant.

We have developed a new method that combines two or more nearly orthogonal scans to generate a model with higher resolution than models created from either of the scans alone. The two scans are first registered to each other and then a net of linked surface nodes is initialized for each of the scans. The nodes from the two nets are then merged and relaxed, subject to constraints set by the resolution of each scan. The result of this process applied to the two data sets of the top figures is shown in the bottom figure. Information from both data sets has been used to generate a smoother and more accurate model of the anatomical structure.

Background and objectives: Generating 3D models of anatomy from medical image data is important for applications such as surgical simulation, planning, and image-guided surgery. An internal scan typically consists of high-resolution data in the imaging plane and significantly lower resolution between imaging slices. The lack of high-resolution information along the scanning direction causes aliasing or terracing artifacts in anatomical models.

Terracing artifacts can be reduced by increasing the resolution of the scan. However, higher resolution between imaging planes means exposure to higher doses of radiation, longer image acquisition times, and higher costs. In addition, in clinical practice, it is more common to acquire a pair of orthogonal scans and use the information in both scans for diagnosis and planning. Our goal is to use the combined information from the orthogonal scans used clinically to generate 3D models with higher resolution than could be obtained from a single scan.

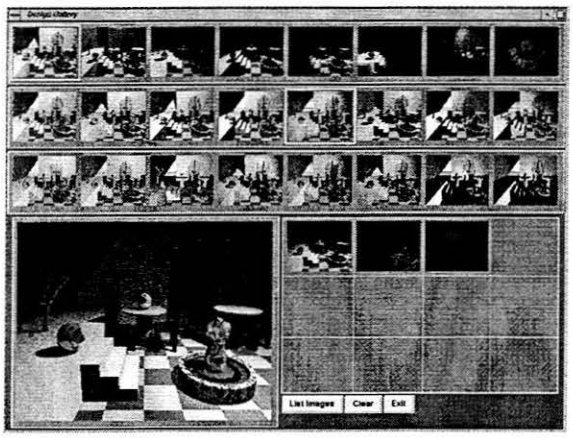
Technical discussion: Dual SurfaceNets extend the original SurfaceNet approach by combining information from two orthogonal volume image scans. The use of Dual SurfaceNets requires a number of pre-processing steps. First, the object of interest is segmented from each of the scans. The segmented images are then registered into a common coordinate frame, and a SurfaceNet is initialized for each of the data sets. We have used a straightforward registration algorithm that aligns the data sets using a stochastic gradient descent on the sum of squared differences of low-pass filtered versions of the segmented images.

The dual relaxation process smoothes the shapes of both SurfaceNets subject to constraints based on the resolution of the two scans. The nets are alternatively relaxed independently, to attain smooth models, and towards one another, so that information from both scans influences the model shape. During both relaxations, nodes are constrained to lie within a volume limited by the original scanning resolutions. The constraints are eased as the relaxation proceeds to reduce the effects of imaging, segmentation, and registration errors. While this easing of constraints means that the resultant model is no longer guaranteed to lie within one voxel from the original segmentations, the final model is guaranteed to lie between the two initial models.

Future Directions: We are continuing to explore better registration techniques, different metrics for adjusting node positions during the dual relaxation step, and methods that will accommodate models with different topologies.

Authors: Mike Leventon (MERL summer intern), Sarah Frisken Gibson
June 6, 1999
http://www.merl.com/projects/surgical/dualSurfaceNets

A Design-Gallery Approach to Lighting Design



See Color Plate #3

DG: Design Gallery

Just as in photography and cinematography, good lighting is crucial for high-quality graphics. However, unlike in the real world, it is still not possible to manipulate virtual lights interactively and see their true effects in real time, because high-quality rendering algorithms are very time consuming. 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 perceptually different images that can be produced by varying the parameters (e.g., type, location, direction, falloff characteristics, etc.) of a single light. This selection is generated as a batch process, without human intervention. Upon completion of this process, the user can then select individual lights from the automatically generated gallery and combine their effects in real time via image addition.

Background and objectives: At the coarsest level of abstraction, all computer-graphics processes map input parameters to output values. For example, a rendering process maps the lighting parameters for a particular 3D scene model to output pixel values in a corresponding 2D image. 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: The principal technical challenges posed by the DG approach are dispersion (finding a set of lighting-parameter vectors that optimally disperses the resulting images) and arrangement (arranging the resulting images for easy browsing by the user). For dispersion we use a form of evolutionary computation, and for arrangement we use a hierarchical interface organized by graph partitioning. Technical details are described in our 1997 SIGGRAPH conference paper.

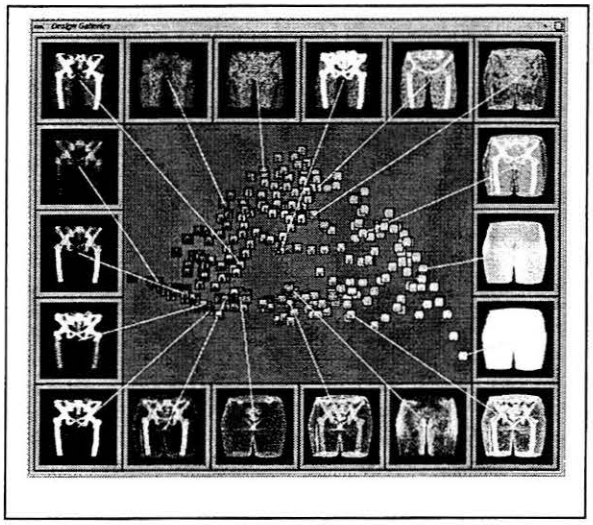
Collaboration: Initial work on the DG/Lighting concept involved researchers from MERL, Harvard University, and the Georgia Institute of Technology. The development of a commercial version of the DG/Lighting software is being undertaken at MERL only.

Future Directions: We are currently developing 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 will be sold through Digimation, a software publishing house that specializes in 3D StudioMax plug-in software. The projected price is between $100-$200, and potential sales range from 500-5,000 copies.

Authors: David Anderson, Joe Marks, Egon Pasztor, Sean Rosser.

http://www.merl.com/projects/design/index.html June 8, 1999

VolDG - Design Galleries for Volume Graphics



See Color Plate #4

DG: Design Galleries.

Direct volume rendering is a key technology for the visualization of large 3D datasets from scientific or medical applications. Of particular importance to the look of direct volume-rendered images are transfer functions. A transfer function assigns values for optical properties, such as color and opacity, to original values of the dataset being visualized. Unfortunately, exploring different transfer functions can be a tedious task, most commonly achieved by ``trial and error."

Managing and organizing the exploration of transfer function space is usually the responsibility of the user; the computer is used as a passive instrument. In the Design Gallery (DG) approach the parameter-setting task is divided more equitably between user and computer. DG interfaces present the user with the broadest selection - automatically generated and organized - of perceptually different images that can be produced by varying a given set of transfer functions.

Background and objectives: At the coarsest level of abstraction, all computer-graphics processes map input parameters to output values. For example, a volume rendering process maps visualization parameters to output pixel values. The trick in producing compelling 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.

Inverse design is a general paradigm for computer-aided design of graphics. Instead of tweaking input parameters, the user supplies an objective function over the output values. This function will generate a high score (say) for good output values, and a low score for bad output values. The computer then searches for a set of input parameters that will maximize the objective-function score. This general approach has been applied to many problems in computer graphics, 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 graphic or animation.

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 input parameters that spans the space of output values 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: The principal technical challenges posed by the DG approach are dispersion (finding a set of input-parameter vectors that optimally disperses the resulting 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 have experimented with hierarchical interfaces arranged by graph partitioning, and browsing interfaces arranged by multidimensional scaling. Details of our technical approach are described in our 1997 SIGGRAPH conference paper.

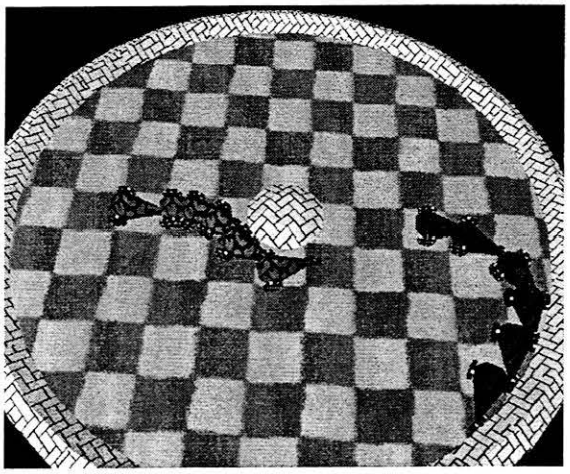
Collaboration: This project currently involves researchers from MERL, Mississippi State University, and the University of Virginia. Initial work on the DG concept involved researchers from MERL, Harvard University, and the Georgia Institute of Technology.

Future Directions: We are currently developing a DG front-end for the popular Visualization Toolkit (VTK). We plan on distributing this code freely to researchers, and to use it ourselves to conduct research in volume visualization. In a related project, we are developing a commercial DG application for lighting design.

Authors: Joe Marks, Hanspeter Pfister
http://www.merl.com/projects/dg

May 18, 1999

Evolutionary Optimization through Simulation



Real-life optimization problems involve complex systems and many variables. Often a designer knows the parameters that can be varied and an objective measure of the quality of a design. However, incomplete knowledge of how the interdependent parameters affect the final quality preclude use of many mathematical optimization techniques. We have explored using genetic algorithms in conjunction with a simulation system to attack such optimization problems. Our focus is on problems that involve machines and mechanisms in the physical world, for example: varying parameters of an assembly line to maximize throughput, varying parameters of a coin-sorting mechanism to maximize robustness, or varying control programs of industrial robots to maximize energy efficiency.

Background and objectives: This research examines the problem of exploring a large parameter space to find high quality solutions to a design problem. Such problems are ubiquitous in all disciplines of engineering. As a sample problem, our method was used to generate effective control programs for two robotic vehicles.

Technical discussion: Genetic algorithms provide a heuristic means of finding good solutions when the relationship between parameters and output is unknown or too complicated to model explicitly. In this case, accurate simulation is needed to determine the output corresponding to a given set of parameters. We tested our algorithms on the problem of learning control programs for two robotic vehicles: a pursuer and an evader. The above figure depicts overlayed time-lapsed images taken during a simulation which pitted particular pursuer and evader designs against one another.

One advance over previous work in the area is the rigid body simulator used to drive the evolution. Our simulator computes the dynamics of three-dimensional rigid-body models, with support for collisions, transient contact forces, Coulomb friction, restitution, and ideal joints. The generality of the simulator revealed nuances in the control strategies that a simpler simulator might have missed. For example, a two-dimensional simulator would not have detected that some control programs caused the vehicles to overturn by turning too sharply.

For a pursuer-evader scenario such as ours, coevolution is technique used to evolve control programs for both species simultaneously, as the ever-improving designs are pitted against one another. This research also compared the effects of testing individual pursers against multiple evader populations that were independently evolved, and vice-versa. Empirical results showed that this style of coevolution produced better designs than coevolution in which single populations of pursuers and evaders were evolved.

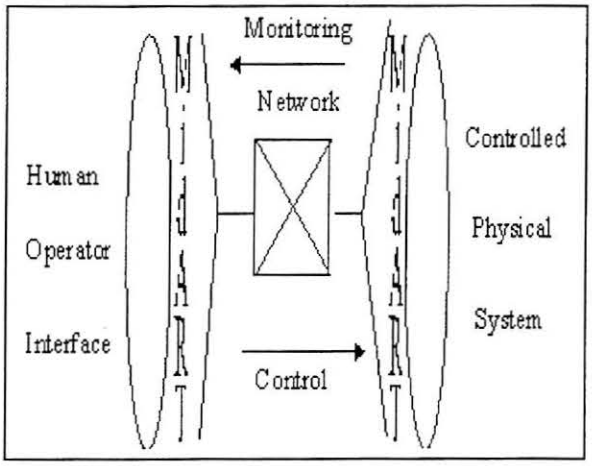
Related to the explorations of large parameter spaces is the Design Gallery methodology. Here, a quality measure is replaced by a similarity measure, and the computer's task is to find solutions that span the range of possibilities.

Collaboration: This research was conducted in conjunction with the Dynamical & Evolutionary Machine Organization (DEMO) Group at Brandeis University.

Future Directions: This method should be tested on real world problems, possibly from the manufacturing domain. A general interface for specifying parameters, quality metrics, and specifics of the genetic algorithm would enable integration with CAD systems.

Authors: Brian Mirtich

MidART: Middleware for Distributed Real-Time Systems



MidART is a distributed real-time software package with easy-to-use programming interface for data acquisition and communication. It allows rapid development of multi-threaded concurrent communicating real-time applications in a network environment. MidART is most useful for applications where humans need to interact, control and monitor instruments and devices through network connections with computer interfaces.

MidART Version 1.0(NT) for Windows NT 4.0 and Version 1.0(POSIX) for Unix platforms have both been released. So far, six 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: It is becoming ever more important to build distributed real-time applications using open, standard, commercially available computers and networks. This is largely due to (1) network and processor technology advances, (2) cost considerations, and (3) the desire for easy system integration and evolution. Today's systems are largely proprietary. Existing network software facilities such as the socket interface is cumbersome and 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 middleware 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 -- 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. These include 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. RT-CRM is not simply a multicast or unicast data transport protocol. It does not restrict the data to be transmitted to the receivers/readers immediately after the data becomes available. In contrast, we provide the application designers the facility to 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. This is accomplished via a set of channel start and stop protocols, and channel bandwidth resource reservation schemes.

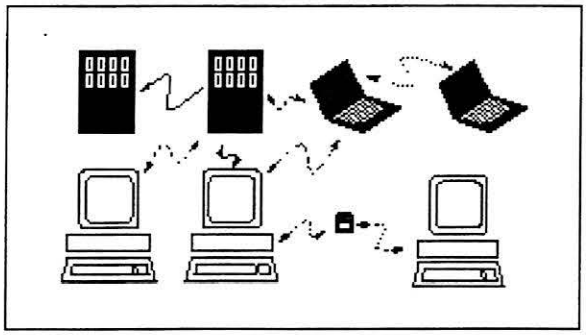
Collaboration: MidART is a collaborative project between MERL and the System Technologies Department in the Industrial Electronics and Systems Lab in MELCO. University of Massachusetts at Amherst has been our academic collaborator.

Future Directions: We would like to apply MidART to a wider range of Internet applications by extending the MidART service model for dynamic Quality of Service Management.

Author: Chia Shen
http://www.merl.com/projects/midart

June 17, 1999

Deployment of Reconcile



After a 9-month beta test, Reconcile has been licensed to an outside company for development and distribution as a product.

Reconcile supports replicating files between multiple computers, specifically:

- Sharing files between office and laptop computers
- Sharing files between office and home computers
- Web site maintenance
- Duplicating an entire file server for online backup

Background and objectives: Home and laptop computers are becoming very common, but are ill-served by traditional distributed file systems, which require continuous connections between clients and servers. Disconnected users need help remembering what files they have changed and which computer has the newest version of each file.

Reconcile automates this by tracking and remembering updates locally and propagating them only occasionally, when the user chooses to do so. Connectivity is required only for reconciliation, and not at other times, providing minimum network demand and dependency on server and network availability. Tracking is accomplished by a unique combination of hash-coding and journaling, resulting in a robust and automatic system which propagates updates safely, detects genuine conflicts, and minimizes false alarms.

Although it performs much the same function as a distributed file system, Reconcile has several unique features:

- Any two sites can reconcile with each other. There is no central file server, and no single point of failure.
- Files from any combination of applications are supported. Reconcile does not use special knowledge of specific application database formats.
- Reconcile does not modify the operating system or local file system, and introduces no overhead during normal operation. It does all its work "just in time" when the user requests a reconciliation.

Technical discussion: Reconcile has been deployed to an internal user community since July 1998, supporting Unix (Digital OSF1, Linux, IRIX, and SunOS) and Windows (95, 98, NT.) Applications involved both individual users and system maintenance:

- About 10 individuals use Reconcile for their working files on home computers and laptops, involving a mixture of Windows NT, 95, and 98 machines, Linux, and MERL's Unix-based central file server.
- MERL's system administration staff uses Reconcile for maintenance purposes. All working files for both MERL and VGO are replicated daily to "hot backup" file servers for use if case the primary file server fails. MERL's web site is replicated from a user-modifiable staging area to a protected "golden" area before being published to the outside world.

User feedback was obtained via an electronic mail list, supplemented by informal personal interviews and a questionnaire. Generally feedback has been quite positive once people get going. An example of a major change resulting from user feedback was a complete revision of the graphical user interface to make the process of setting up new sites more transparent.

In May 1999 we concluded a non-exclusive licensing agreement with NovaStor, Inc, a North American software company specializing in storage management products, to develop and market Reconcile as a product. NovaStore expressed great satisfaction with the quality of the program and its readiness for further development, and are proceeding on an aggressive schedule. They plan to release a product this summer.

Collaboration: Internal users at MERL and the Volume Graphics Organization cooperated by running the program and providing feedback.

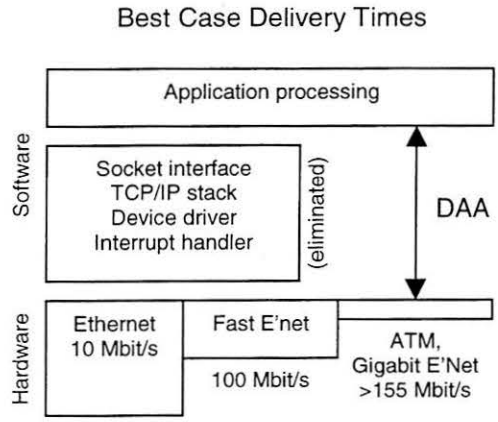
Future Directions: Although the NovaStor agreement constitutes the primary future development path for Reconcile, ITA and MELCO retain ownership of all intellectual property and are free to develop or license it elsewhere.

Authors: John Howard

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

June 14, 1999

Performance of Direct Access Architecture



Direct access allows an I/O adapter and application program to communicate directly through application virtual memory. It can eliminate significant operating system overhead and increase I/O performance dramatically. This is particularly important for network I/O, in which hardware speed has increased dramatically while software overhead has not, but potentially it applies to any form of I/O.

We evaluated the performance of a direct access architecture using the DART Network Interface chip in a PC running Windows NT 4.0. DART achieves an application to application latency of 18 microseconds and a throughput of 95% of the network maximum. This proves both the feasibility and the effectiveness of the DAA approach.

Background and objectives: DART was designed to provide extremely fast latencies and high throughput by supporting direct communication with the adapter through application virtual memory. The goal of this project was to evaluate the function of a DART-based adapter in a real system, to demonstrate a real direct access driver support using Windows NT, and to measure the performance actually attained.

Technical discussion: Direct Access Architecture has several essential parts. The adapter understands virtual addresses using an address translation table. I/O requests and buffers are communicated through the shared virtual memory using ring queues. I/O operations must be associated with the correct application and address space; in the case of DART this is done by establishing virtual connections.

We developed and debugged a network interface card, drivers, libraries, and application programs for DART, and used them to measure performance. Key results were:

- Application-to-application latency for short frames was about 17.7 microseconds, broken down into 4.8 to process the send request, 3.3 to construct the physical representation of the frame, 7.2 physical framing and transmission time, and 2.4 to receive the cell and notify the application. The "raw cell" mechanism, which bypasses the transmit ring queues, saved 2.3 microseconds for a latency of 15.4.
- Latency increased linearly with frame size at a rate of 2.852 usec per cell, close to the theoretical time of 2.83 usec, and for large frames throughput approached 95% of the hardware maximum.
- DART interrupts the host CPU to handle some overhead functions, including address translation faults and rate management cell processing. On a 233 MHz Pentium II system, the time to process such a host service request in the lowest-level interrupt handler was 8.5 microseconds.

Software developed for this evaluation project included a Windows NT 4.0 kernel-mode driver, an application-level library to communicate with it, and several test and measurement application programs. The main function of the driver was to set up the DAA address translation and various other control tables, and to handle host support interrupts. Once the adapter is initialized, the application and adapter communicate directly without driver involvement.

We were also able to map the adapter's control registers directly into the application address space. This allowed us to debug the adapter and develop almost entirely at the application level, using only a minimal driver. If the device register numbering were re-organized to separate "safe" registers such as virtual addresses and most commands from "unsafe" registers like physical addresses, we could exploit this additional direct access capability to perform most of the setup and some additional ring queue processing without kernel involvement.

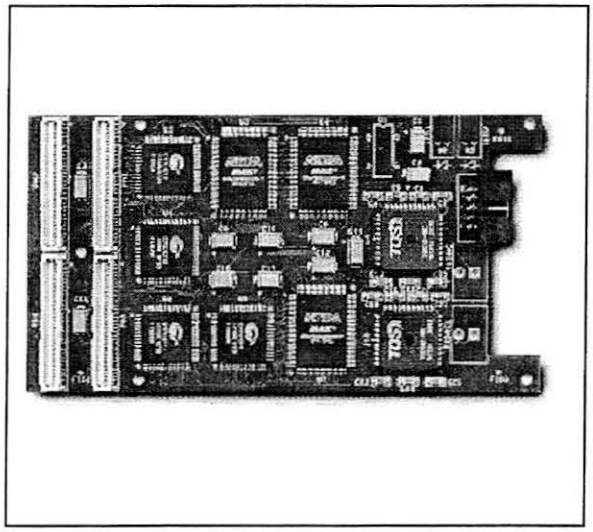
Collaboration: This is a joint project of MERL and Vsis, Inc.

Future Directions: Evaluation boards, software, and documentation are available as needed.

Author: John Howard

http://www.merl.com/projects/dart June 16, 1999

ChaosLAN: A Gigabit LAN Based on Chaotic Routing



ChaosLAN is a network designed for the most demanding requirements of high throughput and low latency in a clustered computing environment. Many applications such as a distributed haptics renderer simply cannot work using a standard network such as Ethernet because the end-to-end network latencies are several orders of magnitude too large. ChaosLAN delivers latencies on the order of a few microseconds. Its mesh-based routing strategy sustains up to 90 percent of the theoretical maximum possible injected message traffic, compared with about 50 percent for a competing high performance network such as Myrinet.

On the left is a photograph of the ChaosLAN network interface card (NIC) prototype, the Chaos Mezzanine Card (CMC). The CMC uses four high-speed Cypress FIFO chips for buffering, three Altera FPGAs for control and data encoding, and a pair of TriQuint Fibre Channel chips for transmitting and receiving serial data at 1.25 Gbit/s. The CMC attaches to the PCI Pamette card from Compaq.

ChaosLAN: Chaos Router Local Area Network. NIC: Network Interface Card. CMC: Chaos Mezzanine Card.

Background and objectives: The chaotic routing algorithm has been simulated extensively in software. Our goals are to create a low cost physical prototype for performing more extensive tests and measurements of the routing algorithm, and transition into production and deployment of this technology. The ChaosLAN NIC (shown above) has been constructed and is fully functional. We also designed and built a mesh router chip to be used in a 16 node router switch. After the router switch is built, we can then complete the evaluation of the entire system.

Technical discussion: ChaosLAN leverages many years of research into networks and routing algorithms at the University of Washington. Chaotic routing is a *non-minimal, adaptive* routing algorithm for 2-D and 3-D mesh and torus routers. The standard packet routing algorithm, used in Myrinet and the Intel Paragon routers, is *non-adaptive* and *oblivious*. The oblivious algorithm cannot avoid hot-spots or congestion in the network. As a result, the measured throughput in a heavily loaded network is only about half of the theoretical maximum that the interconnection wires can provide. Chaotic routing greatly improves the throughput under heavy loads by using randomness to dissipate the congestion at hot-spots. Chaotic routing has been proven to be both deadlock-free and livelock-free, an essential property for making a robust network. These properties are not guaranteed for adaptive routing algorithms in general.

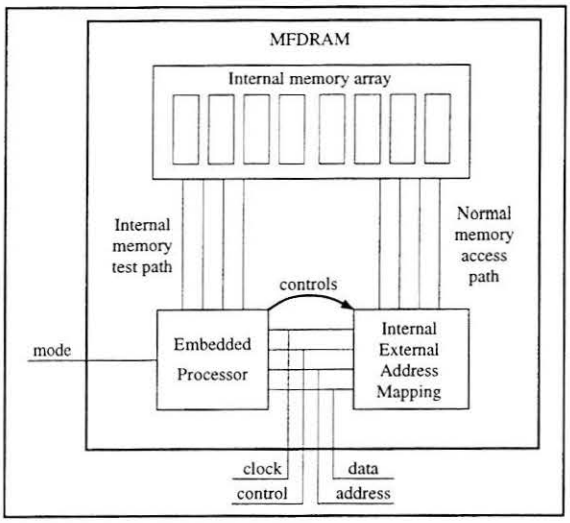
The network interface circuit provides user programs with very low latency access. The implementation of the CMC is based on the Cranium Network Interface Architecture, a framework that is compatible with adaptive routing algorithms such as the chaotic routing algorithm. Cranium is similar to (but developed independently from) the DART ATM NIC, also developed at ITA-MERL. Both approaches provide application programs with direct protected access to the network. The CMC outperforms the DART ATM NIC because it avoids the inherent overheads associated with ATM, and it uses a 1000-Mbit link rather than DART's 155-Mbit link. CMC and ChaosLAN allow a much greater amount of customization than is possible with a complex standard such as ATM.

Collaboration: Chaotic Routing Project group at the Department of Computer Science and Engineering, University of Washington, Seattle, Washington USA.

Future Directions: The ChaosLAN research project is seeking collaborations within MELCO so that the technology can be further developed and eventually deployed either internally or as a commercial product.

Author: Neil R. McKenzie

A Multi-Functional Memory Providing Continuous Self Testing Capability



We describe an architecture for a Multi-Functional DRAM (MFDRAM). The MFDRAM is fully backwards compatible with existing DRAM standards such as the Intel PC-100 standard. Capabilities of the MFDRAM include the following:

- Continuous built-in self test
- Constant fill
- Block transfer
- String matching
- Data compression

A patent has been filed in the USA for the architecture. The concepts need to be developed further so that they can be eventually applied to Mitsubishi Electric's future DRAM product line.

DRAM: Dynamic Random Access Memory. eRAM: Embedded RAM. MFDRAM: Multi-Function DRAM. BIST: Built-in Self Test. IRAM: Intelligent RAM.

Background and objectives: Mitsubishi Electric is a leader in memory chip manufacturing using eRAM. We are seeking new ways to exploit this capability and extend MELCO's leadership in this field. In particular we are investigating architectures for multi-functional memories that provide benefit for the largest possible number of computer users. Our memory architecture allows the memory in the semiconductor to test itself continuously, while simultaneously it allows the DRAM to be accessed normally by the host processor with affecting the access time.

Technical discussion: As DRAM sizes become ever larger, they become statisically more prone to failures in the field. While most hard errors are caught during manufacturing test, and hard errors are therefore rare after the DRAM is bundled and sold in a complete system, they are nonetheless potentially catastrophic. Hard errors can be detected in the field using the chip's BIST (built-in self-test). However, most commercial operating systems ignore BIST entirely. Our idea is to have the chip to execute BIST automatically and continuously, so that the operating system does not need to explicitly trigger it. Redundant copies of some data blocks inside the chip are kept so that the host processor and the internal (BIST) processor do not conflict. Copies are made by the internal processor using an atomic update so that data is never lost. Furthermore, by adding a small amount of intelligence to the operating system, other functions can be executed in the processing element(s) in the MFDRAM chip. These functions include constant fill, block transfer, string matching and data compression. We feel that these functions are the most universal and would provide the most benefit to users of home computers.

Prior approaches to constructing chips that combine DRAM and logic cannot be easily adapted to meet the goals of this project. Examples include university research projects such as the IRAM from the University of California at Berkeley, and Active Page DRAM from the University of California at Davis. IRAM is a forward-looking design that is not compatible with today's personal computers. Active Page DRAM provides some compatibility, but it uses a parallel programming model that is difficult to adapt to current technology.

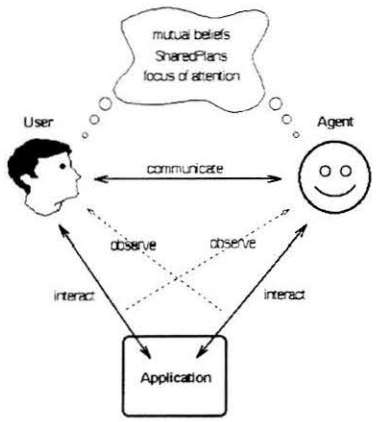
Future Directions: We seek collaborations with semiconductor design groups within MELCO in order to apply the ideas of this research project to future DRAM products.

Author: Neil R. McKenzie

http://www.merl.com/projects/cstdram

April 7, 1999

COLLAGEN: Java Middleware for Collaborative Interface Agents



Computer users want the software they interact with to be more than just a set of tools; they want 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. The software agent can both communicate with and observe the actions of the user, and vice versa.

COLLAGEN is middleware for developing collaborative interface agents. Software developers can use COLLAGEN to implement a collaborative interface agent for any Java application. Release 1.0 of COLLAGEN was delivered to MELCO IESL and ATC in November, 1998.

COLLAGEN (tm): COLLaborative AGENts, GUI: Graphical User Interface.

Background and objectives: COLLAGEN is inspired by the study of naturally occuring human collaboration, such as two people assembling a complex mechanical object or two computer users working on a spreadsheet together. In particular, COLLAGEN implements a computational theory of human collaborative discourse that has been empirically validated across a range of problem-solving tasks.

Technical discussion: A key feature of COLLAGEN is that it automatically maintains a history (log) of the user's and agent's activities. This *segmented interaction history* is hierarchically organized according to the structure of the user's tasks. It helps re-orient the user when he gets confused or after an extended absence. It also supports transformations, such as returning to earlier points in the problem solving process.

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

Although COLLAGEN can serve as the discourse component of a natural language (speech) understanding system for communicating with an interface agent, it is currently being used without natural language understanding. In this configuration, COLLAGEN presents the user with a menu of possible utterances that are predicted by the discourse theory from the current discourse state.

COLLAGEN is implemented using Java Beans (tm) technology, which provides a high degree of modularity and portability between different computing platforms.

Collaboration: The basic theory and architecture underlying COLLAGEN was developed jointly with Lotus Development Corporation (IBM). We are now working with the Power and Plant Systems Dept. of IESL (Sanken) to incorporate Collagen into DOOARS, the new Java version of their GhostHouse GUI-building tool. We are also working with the Information Systems Dept. of ATC (SentanSoken) to build a COLLAGEN-based interface agent for distributed resource allocation tasks, such as construction materials recycling and power trading.

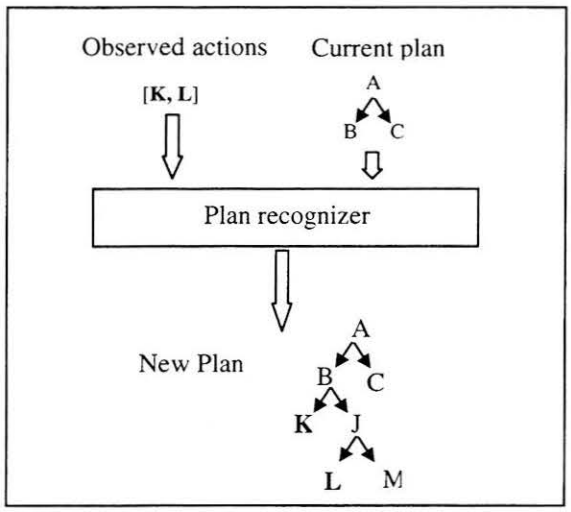
Future Directions: Successful incorporation of COLLAGEN technology into MELCO products, such as plant equipment monitoring, maintenance, and training systems, could provide a competitive advantage by making the complex GUI's in these systems easier to learn and use. COLLAGEN could also be marketed as a separate software product; this would require significant additional investment in productization, including testing, documentation, and other enhancements. We are also continuing work on a number of unsolved fundamental research problems.

Authors: Charles Rich

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

December 1, 1998

Using Plan Recognition in Human-Computer Collaboration



In order for a collaborative agent to assist a user, the agent must know something about the user's goals. We have developed a plan recognition component which infers users' goals from their actions in the context of an ongoing collaboration between a human user and a computer agent. Adding plan recognition to human-computer collaboration significantly reduces the need for users to explictly explain their goals to the agent, allowing for much more natural and efficient interaction.

Our plan recognition component has been included in Release 1.1 of COLLAGEN, a Java middleware system for building collaborative interface agents.

COLLAGEN (tm): COLLaborative AGENts

Background and objectives: Release 1.0 of COLLAGEN was delivered to MELCO IESL and ATC in November, 1998. Release 1.1 of COLLAGEN includes a plan recognition component. Software developers can use COLLAGEN to implement a collaborative interface agent for any Java application.

Technical discussion: An important trend in recent work on interactive computer systems has been to view human-computer interaction as a kind of collaboration. In this approach, the human user and the computer (often personified as an ``agent'') coordinate their actions toward achieving shared goals. A common setting for collaboration, which is the focus of this work, is when two participants can both communicate with each other and observe each other's actions on some shared artifact.

Successful collaboration requires that each participant maintains a set of beliefs about the mutually believed goals and actions to be performed, and about the mutually believed capabilities, intentions, and commitments of the participants (this is part of what Grosz and Sidner called the *SharedPlan*). Each participant also knows a set of methods, called *recipes*, for decomposing goals into subgoals. The role of plan recognition in this framework is as follows: Suppose one participant, e.g., the software agent, observes another participant, e.g., the user, perform an action A. The agent invokes plan recognition to determine the set of possible extensions to its current plan which are consistent with its recipe knowledge and include the user performing A. If there is exactly one possible such extension, the agent adopts this extension as its new plan; otherwise, it may ask a clarification question.

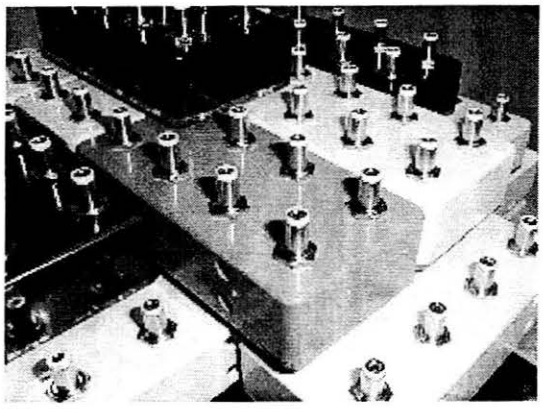
Although plan recognition is a well-known feature of human collaboration, it has proven difficult to incorporate into practical human-computer collaboration systems due to its inherent intractability in the general case. Our investigation, however, has shown that we can exploit three properties of the collaborative setting in order to make plan recognition practical. The first property we exploit is the focus of attention. We can limit the search required by plan recognition based on the observation that consecutive utterances or actions will refer to the same aspect of the plan, unless a shift of the focus of attention is signaled. The second property we exploit is that, in contrast to the general case for plan recognition, we do not need to infer the user's entire plan. Instead, during collaboration, the plan recognizer need only incrementally elaborate a hierarchical description of the current plan to account for new observations. The third property we exploit is the ability of the agent to occasionally ask users questions about their plans and goals.

Collaboration: This research has been done jointly with Lotus Development Corporation (IBM).

Future Directions: Adding plan recognition to COLLAGEN has opened up several new areas of research, including how to participate in a clarification dialogue, how to interpret rejection, and how agents should structure their actions so as to be easily understandable.

Authors: Neal Lesh, Charles Rich

Self-Describing Building Blocks

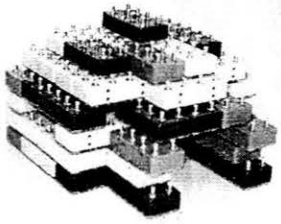


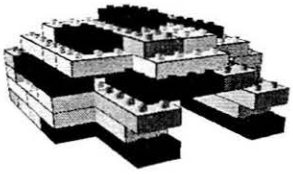
Few people know how to use graphics modeling packages, but everyone can build objects out of blocks. We have developed a set of building blocks that can describe the geometric structures into which they are assembled.

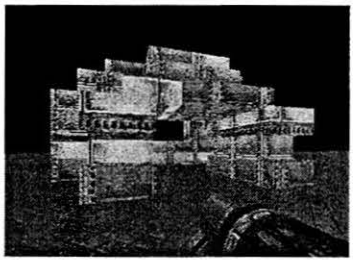
Each building block contains a PIC microprocessor and simple sensors and transducers. The blocks in an assembled structure use a distributed algorithm to discover how they are connected to their immediate neighbors. This information is then relayed from block to block until it reaches a host computer. The structure can then be rendered in various styles, including decorative interpretations in which structural elements are identified automatically and augmented appropriately. Once rendered, the virtual models support a variety of interaction techniques, using the sensors and lights in the blocks.

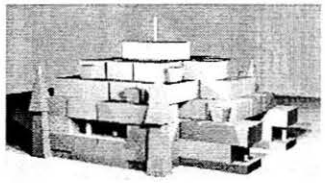
Background and objectives: Possible applications for this technology range from game playing to architectural studies, engineering simulations, and building maintenance.

Technical discussion: Below is shown a physical structure comprising 50 of the self-describing blocks, a literal rendering of the virtual model recovered from that physical block structure, and the same model being used as the setting for a popular game (Id Software's™ Quake II™).









At left the same model is shown being automatically rendered in two predefined architectural styles. 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.

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

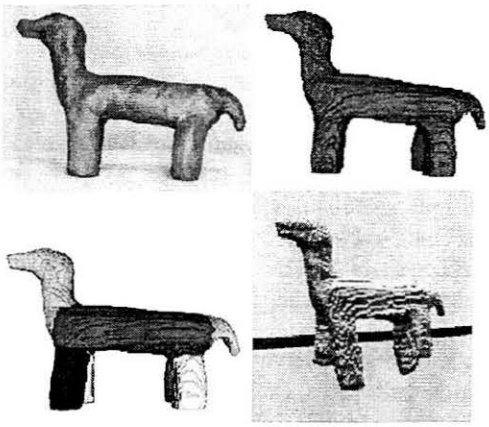
Future Directions: We are exploring a number of new interaction metaphors made possible by this technology.

Authors: David Anderson, Jamie Frankel, Joe Marks, Darren Leigh, Eddie Sullivan, Jonathan Yedidia.

http://www.merl.com/projects/blocks

June 3, 1999

Bringing Clay Models to Life



Creating animated, 3-dimensional characters is difficult. Our goal has been to make this activity accessible to everyone.

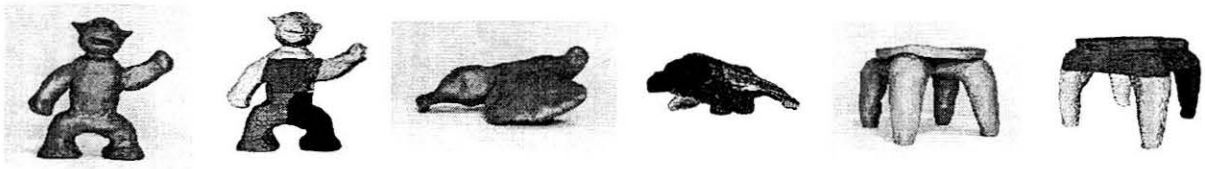
Using modeling clay, we made a variety of models characteristic of what a child might create: people, animals, trees, cars, boats, houses, etc. We captured a volumetric scan of each model, using a sequence of images taken with consumer-grade digital camera and a turntable. The models were recognized, using a small library of object templates, and their component parts identified; in the figure at left, the voxels comprising the dog's head, legs and tail are shown in different colors.

Having identified the different parts of a model, we can animate it appropriately in a virtual world.

See Color Plate #5

Background and objectives: The 3D scanners on the market today are very expensive, and produce very precise 3D models. However, it is now practical to build very inexpensive 3D scanners, having certain limitations. Here we explored an entertaining application that a consumer-grade scanner might make possible. Our application was inspired by the ancient myth of Pygmalion, whose sculpture of a woman was brought to life by Venus.

Technical discussion: The technical novelty in our system lies in our approach to model recognition and parsing, both of which are accomplished by comparing a set of parameterized object templates to a scanned clay model. The templates are deformed to match the model, and the matching score determines how the model is classified. Matching is performed against a total of 130 object templates, 10 from each of 13 categories (biped, quadraped, insect, chair, table, car, boat, bicycle, flyer, rocket, tree, house, and bridge). Once the best-matching template has been found for a given clay model, the match is used to parse the model into its constituent parts. For example, if a model is recognized as a biped, the match between the clay model and the biped template is used to identify the model voxels that constitute the head, arms, legs, and torso. Results from a few models are shown below.



We achieved our goal of bringing a clay model to (virtual) life automatically. The body measurements, masses, and moments of inertia were computed from the parse of one of the quadruped clay models (shown above). These values were passed as input to an adaptive control and simulation system for a four-legged robot, which adapted an existing control system to the dynamics of this particular clay model. The motion data computed by the simulation were then used to animate the object template, which in turn was used to animate the scanned volume by moving its voxels in concert with the linkages of the template to which they were assigned in the parsing phase.

Collaboration: The GVU Center at Georgia Tech worked with us to create the animations.

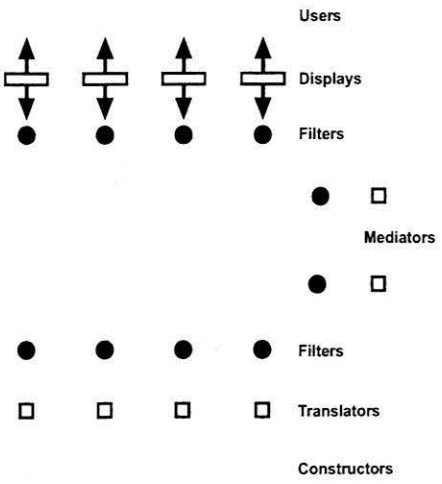
Future Directions: We are working on a better technique for deforming the volume models, so that the animations will appear more natural.

Authors: David Anderson, Aseem Agarwala, Paul Beardsley, Joe Marks

http://www.merl.com/projects/clay

May 27, 1999

Filter Mediated Design: Generating Coherence in Collaborative Design



FMD is an electronic mediating device that allows users engaged in collaborative architectural design to interact with computational agents to generate coherence out of local, domain specific contributions. Agents and users adopt different strategies for producing solutions and interacting with each other, facilitating the creative process. FDM proposes a method for negotiating architectural design across domains, by examining issues of ontology, perception, generation and evaluation. A prototype has been built in which these mechanisms are embodied using computational agents for achieving coherence in a " toy 3D building game" in remote collaborative design. The prototype is implemented on top of ITA's "Open Community", a platform for sharing data over the internet. FMD provides an "object to think with" to explore the dynamics and strategies at play in constructing intelligent designs and design intelligence.

DVE: Distribute Virtual Environment, FMD: Filter Mediated Design

Background and objectives: Architectural design involves the integration of diverse, sometimes conflicting, concepts and requirements into a coherent single composition. The emergence of DVEs, such as ITA's Open Community, provide an opportunity to rethink the ways in which teams have traditionally collaborated, and to offer the mediations needed to augment group negotiation, and foster collective intelligence. Researchers from the fields of Architecture, Artificial Intelligence, and Cognitive Psychology are investigating how DVEs can enhance group design, and how the requirements of collaborative design may help shape the requirements of DVEs. Recent developments in artificial life and human learning suggest that the process of generating coherence, or coordinating multiple perspectives, is decentralized, dialogic in nature, and analogous within and across individuals.

Technical discussion: In both collaborative learning and design, coherence is achieved through iterative cycles of productions and evaluations. At each step, generations and tests are reconsidered in light of each other, and gauged according to some fitness functions. In ill-defined, loosely constrained tasks such as the ones studied here, the search space itself evolves as a function of the form under consideration. The task is to find a representation of the data to be negotiated, and means for accessing and manipulating this shared data, in ways that are sensitive to the dynamics of design. The project has a common geometric and topological database, from which multiple semantic models are constructed by Filters. These semantic models are then negotiated and changes are proposed through the central, semantically sparse database. Coherence emerges when multiple semantic models find satisfaction in the artifact. The project also contributes to the research on Computer Human Interaction. The mechanism supports the distribution of the cognitive load, allowing the computer and human users to focus on particular aspects of the design process. One scenario allows users to freely generate, relegating to the system the task of keeping track of the constraints. In another scenario, users may let the system generate alternatives, exploiting the human ability to appreciate as a lever to sustain the creative process.

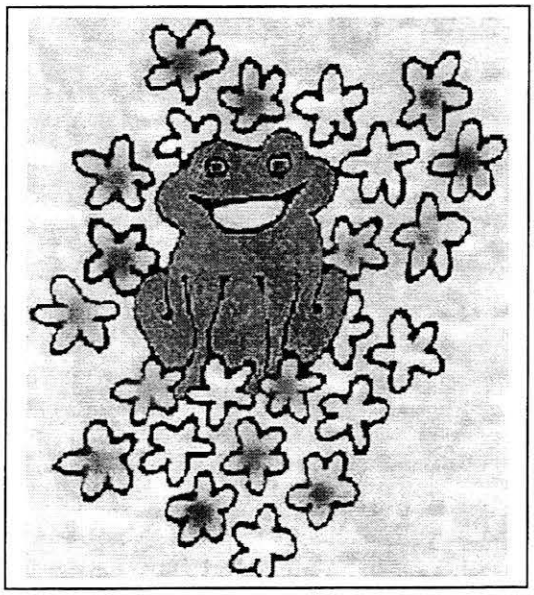
Collaboration: Exploring the potentials of ITA's "Open Community", this project involves extensive input from the DVE team at HSL. The Space Planning and Organization Research Group (SPORG) in the MIT Architecture Department provides the Architectural Experience in Collaborative Design.

Future Directions: The project provides a framework for thinking about the processes of generating coherence in design. Areas of future work can be found in the ontological representation of the world, in the 'filter', 'constructor', and 'mediator' mechanisms which account for the major processes of design, and in the communication between these mechanisms.

Author: Edith Ackermann

June 29, 1999

Experience Journals: A Web-based Tool for Sharing Stories



Researchers at MERL and Boston's Children's Hospital have collaborated to produce an application that records, organizes, and displays stories. The stories are written by people who share problems or interests, such as coping with the illness of a loved one. Characteristics of these communities and their recorded experiences have led us to experiment with an initial algorithm toward the vision of a self-organizing, self-evolving, web-based system.

The algorithm relies on similarity scores and multidimensional scaling to display similar stories near one another. Proximity, rather than coordinate location, is the measure of similarity. Thus clusters of similar entries form naturally. Color coding further indicates how recently stories have been entered in the database and how recently they have been accessed

With support from the Noonan Fund, psychologists at Children's Hospital are conducting studies to ascertain the software's usability and its usefulness in helping people cope with illness.

EJ: Experience Journals

Background and objectives: Often, chronically ill children and their families face illness and hospitalizations without intervention to ease emotional stresses. People who have experienced similar medical problems could offer support, but communication between such families is rare. We are investigating the use of networked computer technology to facilitate this process. With the EJ software, members of a medical community, or any group with common interests, can record text and multimedia vignettes that are made available for browsing at a secure World-Wide-Web site. Because informal communities typically cannot support an administrator to organize and monitor a rapidly growing Web site with large, distributed authorship, the algorithm automatically organizes and manages the entries. Ease of use is particularly important for such varied populations.

Technical discussion: Our algorithm computes a word vector for each EJ entry by removing stop words (e.g., "the," "and," "is," etc.), stripping suffixes (e.g., "ly," "ing," etc.), and weighting the remaining word stems according to the inverse of their occurrence frequency in a large text corpus (i.e., rare words are weighted heavily, and common words discounted). The algorithm then calculates a similarity score for a pair of entries by taking the dot product of the two associated word vectors. We use the well-known technique of multidimensional scaling to position entry icons so that similar entries are near each other in the visual display, and dissimilar entries are far apart.

Psychologists are conducting clinical trials in which Children's Hospital patients and their families use the software to record and share experiences of chronic illness. The researchers are combining observation, interviewing, questionnaire data, and analysis of EJ narratives in order to ascertain the system's therapeutic value and to provide developers with information for improving successive versions of the software.

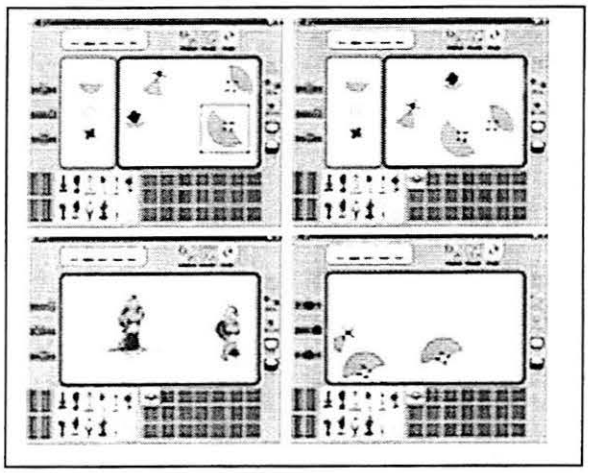
Collaboration: MERL contributors include Dennis Bromley, Carol Strohecker, Joe Marks, Edith Ackermann, Sarah Gibson, Chia Shen, and Marina Umaschi. Contributors from the Department of Psychiatry, Children's Hospital, Harvard Medical School, include Joseph Gonzalez-Heydrich, David Ray DeMaso, Julie Dahlmeier Erickson, Kevin M. Brooks, Beth Donegan, Sarah Lualdi, and Judith Karlin. The Noonan Fund is supporting clinical trials of the software and study of trial results.

Future Directions: We are investigating partnerships for further development of the software, according to results of the clinical trials.

Author: Carol Strohecker

http://www.merl.com/projects/hospital/xpjournal/index.html June 4, 1999

AnimMagix: Learning about Emergent Effects of Behavioral Attributes



See Color Plate #6

AnimMagix is one of a series of prototype environments in which learners experiment with part-whole relationships by composing objects and observing effects that emerge when the objects combine in a larger context. Such experimentation can support development of scientific understandings in the domain of multivariate systems.

In AnimMagix, learners play in a world of whimsical creatures and social behaviors, from which they create groups of dynamic, mutually responsive beings. Interactions take the form of a dialog as learners use different modes of the dynamic interface. Play areas change size to reflect the dialogic turn-taking as constructions proceed.

Users construct creatures by adjusting three behavioral attributes: perceptivity, sociability, and motility. Interactions among these attributes affect the creatures' movements as they "dance" together.

Background and objectives: Color, animation, dynamic modeling, and interactivity are capabilities of current computational media that help in making tools for learning. Researchers take different approaches to the design, implementation, and evaluation of computational learning environments. MERL's approach follows a "learn by doing" philosophy. Like Dewey, Piaget, and Papert, we believe that learners construct their own knowledge, rather than merely absorbing what others present to them. We observe that learning happens particularly well when the learner is making something personally meaningful, which others can appreciate.

Technical discussion: The main significance of the Magix series is in its interaction design. We model interactions as a conversation between the learner and the software toy. In AnimMagix, each of these partners has a specific role in the construction of behaving creatures and groups, which are the parts and wholes of the sociodynamic microworld. The balance of control plays out as working areas on the screen shrink and grow while affording the learner different tools and possibilities for interaction. The learner endows creatures with basic drives and sensorimotor capabilities, and AnimMagix incorporates creatures within the larger context of an interacting social group. The learner can save creatures for play in other modes. AnimMagix includes modes for modifying creatures' drives and behaviors directly or through adjustments to conditions of the environment.

Creatures' interactions are displayed within activation areas modeled on Cartesian coordinates. A creature's user-assigned motility pattern and program-assigned value for effecting changes of location determine where the creature's perceptual cone lies and thus what other creatures it can influence. For each time beat, a creature checks the area within the depth and angled breadth of its cone, to see if any other creatures are present in that area. If the creature (creature A) determines that another creature (creature B) lies within A's perceptual cone, A uses its sociability attribute as the basis for changes to the (x,y) values that establish B's location. A can draw B toward it or send B away. These simple principles yield increasingly complex dynamics as the user adds creatures to the field.

Collaboration: The Magix series is a collaboration between MERL researchers Edith Ackermann and Carol Strohecker, in cooperation with Kazuo Kyuma and Shinji Komori of MELCO. Adrienne Slaughter, Aseem Agarwala, and Daniel Gilman are additional contributors to AnimMagix.

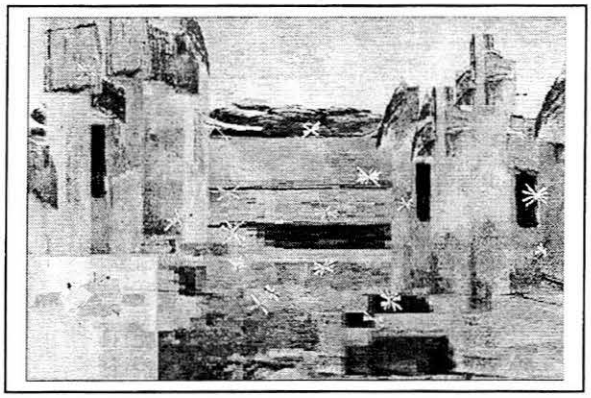
Future Directions: The current prototype is a Java implementation with a point-and-click, mouse-based interface. We are working on designs for a portable platform and a multiple-function stylus input device. In the next versions we plan to recall PatternMagix more directly by displaying two activation areas: one that replicates a construction so that multiple clones can interact, and another that (like the current implementation of AnimMagix) injects a single copy of each new construction into the activation area so that many different kinds of creatures can interact.

Author: Carol Strohecker

http://www.merl.com/projects/amagix/index.html

June 4, 1999

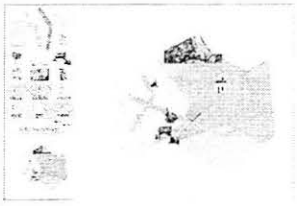
WayMaker: A Design Tool for Virtual Environment Structure

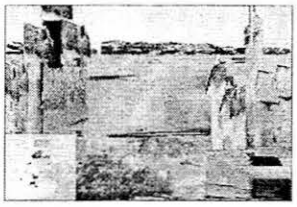


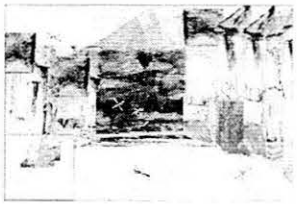
Normally we read maps with the expectation that we will find in the world what we see on the map. What if we designed maps with the expectation that we will create in a world what we place on the map? That is the premise of WayMaker, a design tool based on Kevin Lynch's elements of the "city image." *Districts* are broad regions, such as neighborhoods; *edges* are boundaries between one region and the next; *paths* are the channels along which people move; *nodes* are foci to and from which people travel; and *landmarks* are punctuation points used for general orientation. WayMaker users duplicate, stretch, and position representations of the elements to form a map. The software calculates and displays street-level views along the pathways.

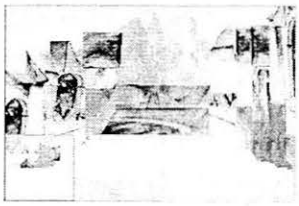
Background and objectives: In text-based multiuser environments, participants interact by socializing with each other and by creating extensions to the virtual environment. They use an associated programming language to construct characters, objects, and spaces. However, this constructive component is not easily transferred to graphical domains. The tools for developing images tend to be complex and hard to use. Yet, graphical virtual environments are emerging as milieus with potential for learning, entertainment, and socializing. Their inhabitants will want to create images of environment components such as avatars, buildings, and landscapes. We need to develop easily usable tools to support such interactions. WayMaker is a tool for designing specifically the structural features of graphical virtual environments. It is meant to complement facilities for designing other attributes of multiuser domains, such as architectural and social features. Because Lynch's elements are essentially topological, WayMaker can also be used as a tool for research in spatial cognition. As WayMaker transforms user-constructed maps, the layouts undergo changes of scale, view, and representation. Users' responses to these changes can help reveal their understandings of basic spatial relationships.

Technical discussion: In this second version of the prototype, we simplified the construction process by eliminating the so-called walkway ribbon. This structure was redundant, as paths already are one of the elements with which users construct maps. Now, views are composited along the paths. Another change is the use of a slightly transparent, miniature version of the map, which is superimposed over the composited views. This device eliminates the need for a double-screen system while maintaining the benefit of simultaneous displays of the map and views. An additional tracking/navigational device, a red dot on the map, locates each view and allows users to indicate where they would like to view the environment. Users can easily reposition the dot to move around the environment, and can easily go back and forth between the map-construction mode and the environment-view mode.









Collaboration: MERL researcher Carol Strohecker developed the initial prototype with Barbara Barros, previously of the MIT Department of Urban Studies and Planning and now of StrataVarious Software. Maribeth Back, of XeroxPARC, is the sound designer. Adrienne Slaughter and Daniel Gilman are contributors to the revised prototype.

Future Directions: In the next version we aim to add sounds and improve the automatic placement of elements within the views.

Author: Carol Strohecker

http://www.merl.com/projects/waymaker/index.html

June 4, 1999

VGO – Volume Graphics Organization

The Volume Graphics Organization (VGO) was created in 1998 to develop leading-edge volume graphics rendering technology. The first product, VolumePro500 (shipped in May, 1999) is a graphics board that does volume visualization. It has a proprietary ASIC chip developed by VGO, which enables real-time volume visualization in real time on a standard PC. VolumePro500 will establish a leadership position in the new field of volume graphics techniques and methods. Over time, VGO integrate its volume graphics technology with traditional 3D polygon graphics.

The VGO team consists of software engineers, hardware ASIC designers and board developers. The team created VGO's first the product in only 18 months from start to first customer ship, an aggressive task for a 13mm die chip with software support.

The **VolumePro500** (code-named Roadrunner) allows PC-class computers to support high-quality, real-time volume rendering. The heart of VolumePro500 is the vg500 chip. This chip can processes 500 million volume samples (voxels) per second supporting gradient estimation and Phong lighting on a per-sample basis.

The VolumePro500 is a PCI board containing one vg500 chip with 160 megabytes of memory. The VolumePro500 integrates with a conventional 3D graphics card for display. A companion CD contains software drivers and a volume library interface (VLI). There are software drivers for Microsoft NT and Sun Solaris systems. The VLI allows access to all the features of the vg500 and provides software support for supervolumes and supersampling in the image plane.

Features of the VolumePro500 include: rendering up to 30 frames/second on 256x256x256 voxel cubes, 32-bit RGBA pixel outputs, interactive parameter changes for modifying: view direction, the number of light sources, Phong lighting, and classification. It can handle an arbitrary cut plane with thickness and box cropping and has a 4K color lookup table, and a 3D line or plane cursor.

The **VolumePro1000** (code named Condor) will be the second board product from VGO and is planned for delivery in 3Q 2000. It will build on the capabilities of VolumePro500 with key enhancements required by users who need to work with volume rendering.

The VolumePro1000 will be a PCI board with up to 500 megabytes of voxel memory and a new ASIC chip, the vg1000. The vg1000 is expected to process 800 million voxels per second. In addition, it will support more flexible voxel formats, including RGBalpha voxels, luminance-alpha voxels, and the ability to set categories on a per-voxel basis. A key advance will be the ability to draw opaque polygons with OpenGL and embed them into the volume.

Engineering design enhancements include the removal of restrictions on the size of a volume that can be rendered at one time, supersampling in x, y, and z in hardware, up to four voxel fields (each separately definable), four separate lookup tables, the ability to bypass lookup tables, and classification and interpolation in either order.

Recent Major Publications

H. Pfister, "Architectures for Real-Time Volume Rendering", *Journal of Future Generation Computer Systems*", Elsevier Science, Vol. 15, pp. 1-9, 1999.

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

B. Lorensen, H. Pfister, C. Silva, and L. Sobierajski, "Architectures for Volume Rendering", *in Advances in Volume Visualization*, A. Kaufman (organizer and speaker), in Proc. SIGGRAPH '98.

Recent Articles About VolumePro

"Volume Rendering Adds Realistic Layers to 3-D Graphics", Anthony Cataldo, *EE Times*, March 10, 1999.

"This Board Speaks Volumes", J. Donlean, *Computer Graphics World*, December 1998.

"VolumePro Takes New Approach to 3D", Peter N. Glaskowsky, *Microprocessor Report*, Vol. 12, No. 15, November 16, 1998.

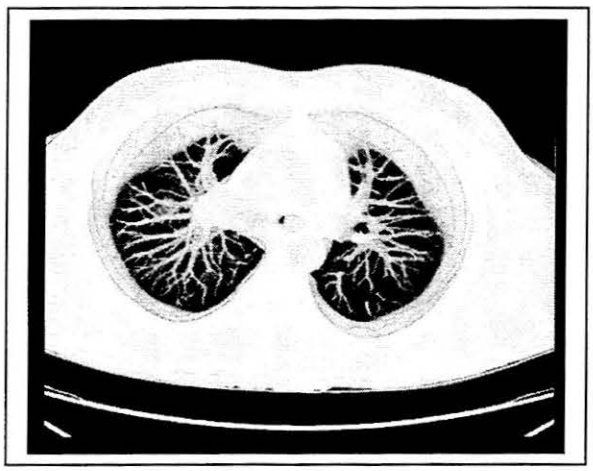
"Applications and Alternatives for 3D Acceleration", Jon Peddie, *The Peddie Report*, Vol. XI, No. 41, October 19, 1998. Article on "Mitsubishi Real-Time Volume Renderer - vg500".

"Accelerators Strain 3-D Graphics Development", Rick Boyd-Merrit, *EE Times*, October 14, 1998.

"Mitsubishi VolumePro: Volume Visualization Hardware for PCs", Roy Latham, *Real Time Graphics*, Vol. 7, No. 4, October/November 1998.

Project Reports – VGO

VolumePro 500 (Code Named Roadrunner)



The Volume Graphics project has been organized to develop leading-edge Real Time Visualization technology and to drive into initial product by May 99. An engineering team has been created to build products to enable compelling visual experiences of unrivaled fidelity by exploiting discrete and geometric representations.

With the first product, VolumePro500, Mitsubishi Electric will deliver real-time volume rendering hardware and software to PC class computers. VolumePro is distinctive for its image quality as well as its performance.

Background and objectives: Starting with a license for SUNY New York technology for rendering volumes, the team invented useful real-time volume rendering technology via a chip, a board, and software. The board fits into a pci slot in a PC.

In January 1998 a plan was put in place which included a functional specification for a first product plus a product roadmap which included 3 subsequent versions of the product to drive the market to use of volume rendering. A business plan was in place and the project was funded and hiring began in earnest. Goal was to have a product for first customer ship in Q1 FY99 that could do real time volume rendering on a PC.

The first functional specification was limited by the engineering understanding of the market. Once marketing feedback of the business opportunity came back, the product was modified to include 6 weeks of additional work to meet customer requirements. A good deal of analysis was then done to determine that the contents of the second chip was what was really needed by the market, and that work commenced in early 99.

In April of 99 the first chip samples came back from the foundry, validating the approach and the design quality and the ability to render volumes in real time with a single chip. This exercise was strong enough to plan for releasing the first products in May with full customer ship in June of 99.

Technical discussion: Volume rendering has matured as a method for visualizing 3D sampled, simulated, and synthetic datasets. Unlike surface-based representations, volumetric representations 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, e.g., for the visualization of a beating heart under real-time 3D ultrasound.

Collaboration: The Real Time Visualization team is a project of the Semiconductor Business Group, funded to create volume graphics product(s). Many applications of this innovative graphics hardware are being discovered in various business units, and it is expected that MELCO will use the board in future versions of radar applications, medical radiation therapy planning devices, in applications where visualization is key. The team also collaborates tightly with MERL in ITA for advanced graphics technology and with ATL and HSL from ITA as needed for broadcast and related technology.

Future Directions: Strong research to take advantage of the knowledge of efficient volume rendering is now a basis for understanding how to handle discrete and geometric representations. General knowledge of the 3D polygon area added weight to the creation of newer methods for handling 3D that can give better image quality and better performance. A product roadmap shows that once version 2 of this product [code named CONDOR] is completed, the next version will have to be able to handle technically challenging areas such as perspective.

Author: Beverly J. Schultz
http://www.3Dvolumegraphics.com

June 29, 1999

Software API, Drivers, Applications for VolumePro500

VP500-1	**Board only**
VP500-1N	**Board with NT driver**
VP500-1S	**Board with Solaris driver**
VP500-2N	**SDK for NT**
VP500-2S	**SDK for Solaris**
VP500-3MN	**Multi-user SDK for NT**
VP500-3MS	**Multi-user SDK for Solaris**
VP500-4	**SDK for other OSs**
VP500-4M	**Multi-user SDK for other**

The Real Time Visualization business has products that are more than just a board. Some of the software is standard with the different products that are available.

The software available for these products is produced in the Volume Graphics group, and incudes an NT driver, a Solaris driver, a Volume Library Interface to allow applications to take full advantage of the advanced graphics hardware, and a Software Development Kit for NT and Solaris. These are standard.

In addition, a software application called VolView is available for sale to researchers and others who want a sample application with source code that they can then modify themselves for development of their own applications.

Background and objectives: Customers of VolumePro500 receive a board and they receive a CD with software. The plan was that this CD would give a full software developers' kit which would have demos, sample applications, a full Volume Library Interface, and drivers sufficient for the user to take full advantage of the VolumePro500 board. This has become a reality.

Technical discussion: Software drivers have been created with full understanding of the need for high performance. They are thorough in interfacing with the hardware and they optimize as needed for both Solaris and NT.

The Volume Library Interface [VLI] has been created to give full use of the hardware to the application developer, and in addition it contains such innovative software as support for supervolumes.

VolView is an application that works with VTK, a freeware software toolkit package, and allows researchers to use that application to understand the use of volumetric rendering. The researcher can take the application sources, also available on the CD, and modify them to create new applications relevant to their own work.

Demonstrations are also available on the CD in the form of sample AVS toolkit applications which give users a good idea of how to use volumetric data, and which give AVS users the opportunity to use the real product updated to handle volumes and to handle the VolumePro500 hardware.

Collaboration: The Real Time Visualization team has collaborated with the AVS company to get demonstrations and with Kitware to get demonstrations and a research application that is being sold with the standard software as a Mitsubishi Electric product.

Future Directions: Software is needed to take full advantage of the VolumePro500 board. At this time, the hardware and its enabling software is being sold as a product. We expect that over time the team will work with customers to ensure that applications will profit from being written to take advantage of the hardware, but that also over time applications that really optimize the use of this hardware will be developed or purchased and available through the Real Time Visualization team.

Author: Beverly J. Schultz

Color Plates

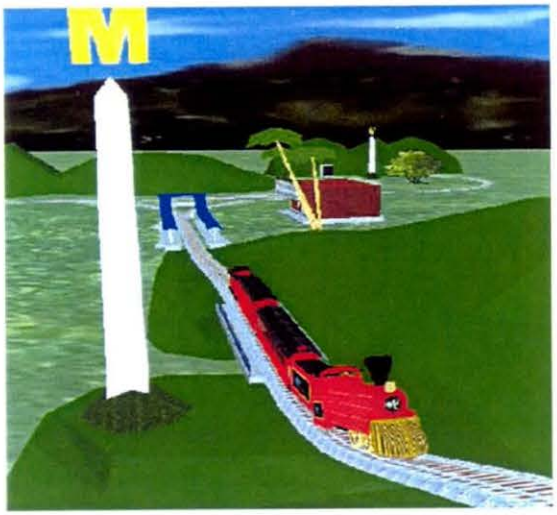


Plate 1 – See page 32

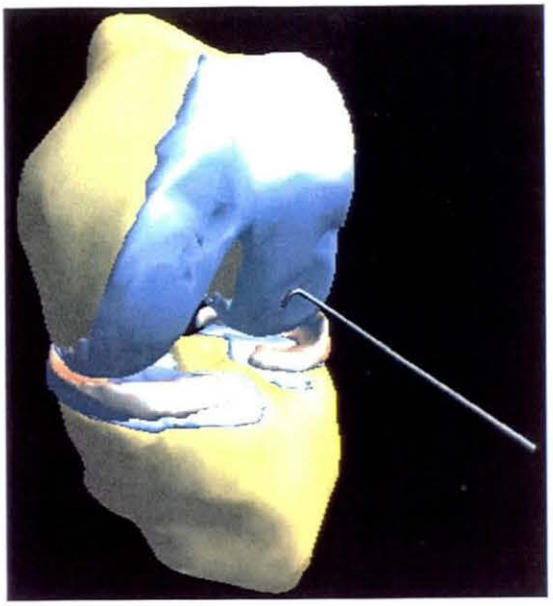


Plate 2 – See page 65

Plate 3 – See page 67

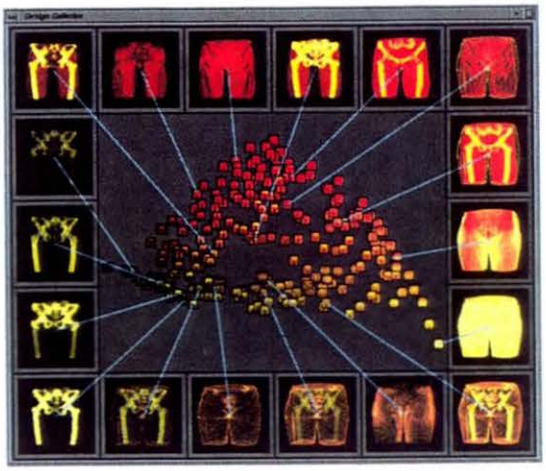


Plate 4 – See page 68

Plate 5 – See page 78

Plate 6 – See page 81