Proposal to:
Department of Transportation

Innovative Railroad Information Displays for Dispatch and Train Engineer Functions

from:
The Media Laboratory
Massachusetts Institute of Technology
20 Ames Street
Cambridge, MA 02139

Principal Investigator:

Ron MacNeil
E15-447
(617) 253-4416
ronmac@media.mit.edu
fax: (617) 258-6264

Proposed Period: 3/1/96 - 11/30/96
Proposal Funding: $140,000

February 29, 1996

Table of Contents

Statement of Work

Innovative Railroad Information Displays for Dispatch and Train Engineer Functions

1. BACKGROUND

A useful information display system must be able to maintain the users sense of context by keeping the important information clearly viewable and conserving important relationships, while subduing information not currently relevant.

The Visible Language Workshop (VLW) of the MIT Media Laboratory has developed unique methods to interconnect different user views of dynamic transportation events by managing spatial information. Some VLW displays can interactively sense and learn about the user's real world context or about events occurring beyond the users environment. VLW has demonstrated that displays can provide alerts or decision support based on digital "expert" knowledge.

Some examples of relevant VLW work to date are presented below:

Large Scale High Resolution Displays:

In 1989, VLW developed what continues to be the worlds highest resolution large scale display, with 6000 by 2000 pixel resolution, using three 2 kilobyte by 2 kilobyte cathode ray tube displays. This system can display ten times as much information as the standard desktop monitor. Thus, an operator can keep the whole dataset viewable while simultaneously viewing multiple local areas of interest using VLW techniques.

Smart Maps:

The GeoSpace Project used activation spreading network technology to develop maps, which had map expertise built right into the lowest level map representation. The map expertise reacted to user requests for a specific type of information. GeoSpaceII applied this technology to the U.S. Census Tiger Database.

Air Traffic Control (ATC) Demonstration:

The ATC system was the VLW's first excursion into fully dynamic 3-space-plus information management. This prototype environment made visible and interactive a concept of future air traffic. In 1995, the system was adopted by Hughes Aviation. According to Dr. Winkler, VP Special Projects for Hughes: "The video and interactive system which we received from the MIT Media Lab and enhanced was well received by the exhibit attendees at the 1995 Air Traffic Controllers Association Conference. We believe this system serves as a good research tool for the evaluation of future Air Traffic Management systems and visualization concepts including Free Flight, Decision Support and Human Centered Automation."

Some information management concepts explored in the above prototype are:

- space between aircraft or greenspace;
- space between aircraft and ground points or altitude lines and range rings;
- virtual air corridors for "look ahead/look behind " context;
- multiple aircraft flight path simulations;
- virtual airport infrastructure;
- multiple interactive viewpoints -- Tower, TRACON, Pilot, Sky-eye;
- weather patterns with "what if" planning options;
- replay capability for preview, training, debriefing, and planning;
- ground/air maps;
- automatic procedure alerts signaling unacceptable aircraft movement;
- multiple viewer 3-D visualization and context control.

2. OBJECTIVE

The objective of this work is to develop and demonstrate a novel railroad information management and display approach which may lead to improved productivity and safety.

3. SCOPE

The VLW will investigate how its methodologies can be adapted to enhancing information management tasks facing dispatchers and train engineers. Assumptions about near term railroad operations, such as use of physical and communication structures or GPS, will be considered. Using software rapid prototyping techniques, one or more concepts will be investigated in VLW's advanced graphics environment and modern PC-based advanced graphic systems.

4. TASKS

Task 4.1 Review of railroad industry information technology.

The contractor shall become acquainted with current information and display technology by literature review, telephone, or personal contacts.

Task 4.2 Observation of railroad operating use of information displays.

The contractor shall observe railroad operating use of information by visits to users and developers.

Task 4.3 Acquisition of information datasets (if needed).

Task 4.4 Rapid prototype of information display software.

Using rapid prototyping software techniques, the contractor shall do a best effort investigation of novel concepts using the VLW's high-end advanced graphics environment and modern PC-based advanced graphic systems.

Task 4.5 Preparation of a videotape.

The contractor shall use its best effort to prepare and deliver a videotape which demonstrates salient aspects of the prototyped systems, displaying simulated railroad operations with information relevant to the dispatch and engineer functions. The specific context and details of this tape will be agreed upon by the USDOT Volpe National Transportation Systems Center (Volpe Center) and the MIT Media Laboratory, Visible Language Workshop.

5. REPORTING REQUIREMENTS

5.1 Letter Reports (Tasks 4.1 and 4.2).

Brief description of relevant contacts and observations

5.2 Informal Progress Reports.

Submitted monthly to the Volpe Center Technical Monitor.

6. DELIVERABLES

6.1	Draft prototype system	Month 4
6.2	Draft Video Script and Informal Storyboard	Month 5
6.3	Draft Final Video Demonstration	Month 6
6.4	Final Video Delivery	Month 8

7. PERIOD OF PERFORMANCE

All work to be completed within nine (9) months after date of award.

8. ENGINEERING COST ESTIMATE

Department of Transportation
Innovative Railroad Information Displays for Dispatch and Train Engineer Functions
March 1, 1996 - November 30, 1996
Principal Investigator - Ron MacNeil

Salaries	**Base**	**%**	**Amount Year One**	**GRAND TOTAL**
Ron MacNeil	$67,200	45%	$22,680	$22,680
Technical Assistant	$26,667	35%	$7,000	$7,000
Administrative Assistant	$28,350	35%	$7,442	$7,442
Research Assistant	$16,572	100%	$12,429	$12,429
Total Salaries			$49,551	$49,551
Employee Benefits (44.5%/45.5%)			$22,325	$22,325
Total Salaries + E.B.			$71,876	$71,876
Operating Costs				
Materials and Services			$3,165	$3,165
Telephones			$50	$50
Xeroxing & Printing			$500	$500
Postage			$100	$100
Travel			$8,000	$8,000
Total Operating Costs			$11,815	$11,815
Modified Total Direct Cost			$83,691	$83,691
Indirect Cost (54.5%/56%)			$46,309	$46,309
Capital Equipment			$10,000	$10,000
Total Costs			$140,000	$140,000

9. BUDGET JUSTIFICATION

Salaries:

Ron MacNeil will provide oversight and supervisory functions (45% effort) during the period of the project.

An administrative assistant (35% effort) and technical assistant (35% effort) will participate with clerical and administrative duties by supporting Mr. MacNeil. Their respective equivalent salaries will be supported from this proposal for the 9 month period.

Employee Benefits:

Benefits are calculated at the approved ONR (proposal) rate of 44.5% through 6/30/96. The rate is then raised to 45.5% from 7/1/96 through 11/30/96 for the duration of the contract.

Operating Costs:

Operating costs represent modest costs for projects of similar scope.

Materials & Services:	For items needed to support Ron MacNeil and his staff while conducting his research on this project, including office supplies, film and developing costs, and other materials and services.
Telephones:	For long distance calls in relation to the research being done for this grant.
Xeroxing & Printing:	For papers that are presented at meetings and conferences. Also for submission of research papers to journals for publication.
Postage:	For all mail associated with the research being done for this proposal, i.e. the mailing of papers for submission to journals; correspondence with sponsor, etc.

Travel:

The travel budget will fund two trips to Washington, D.C., one trip to Ft. Monmouth, N.J., and one trip to the west coast.

Indirect Costs:

Indirect costs are calculated at the approved ONR (proposal) rate of 54.5% through 6/30/96. The rate is then increased to 56% from 7/1/96 through 11/30/96 for the

duration of the contract. The Modified Total Direct Cost Base excludes the cost of capital equipment.

Capital Equipment:

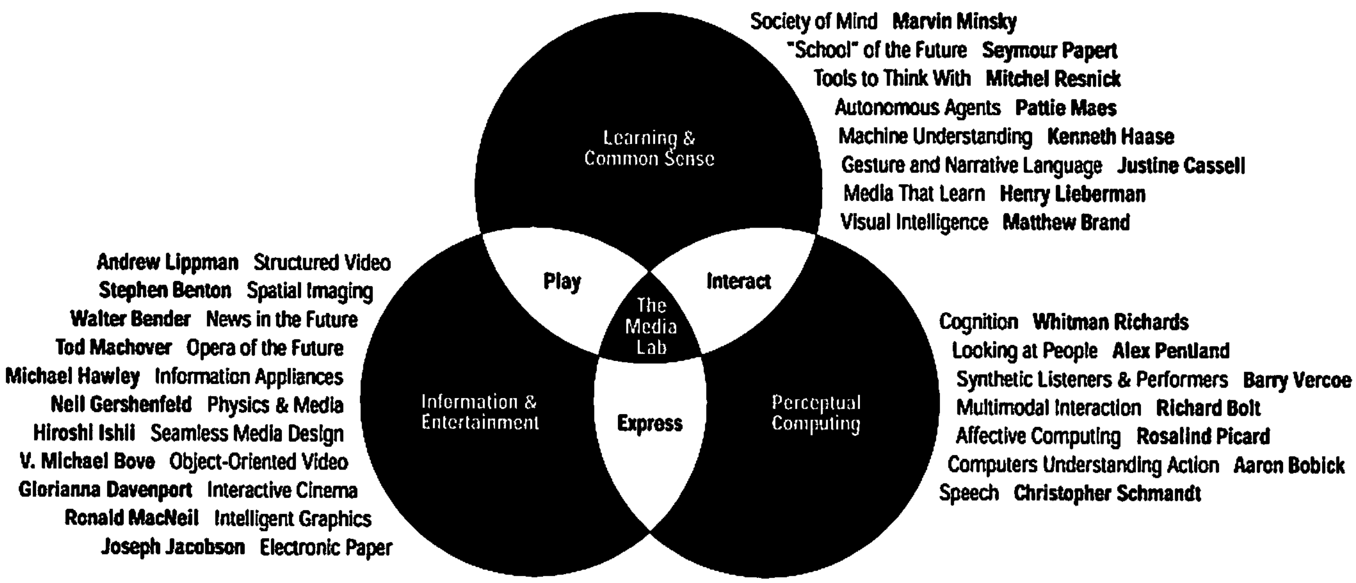
$10,000	For the purchase of a PowerMac, an Apple Newton, and Intergraph boards.

Appendices

Media Laboratory Facilities and Equipment

January 1, 1996

The Media Laboratory is a $50 million facility for advanced study and research in information technologies. The on-going research extends across a wide realm of activities, which are clustered into three broad sections (as illustrated below). These three sections share 110,000 square feet of space, divided into specialized and shared experimental areas to foster both independent and collaborative research. The Laboratory includes 19 faculty; 14 full-time research and academic staff; 46 technical, administrative, and support staff; 124 graduate students and 69 undergraduate researchers. It has a total research volume in excess of $15M per year.



Shared Facilities

- Philippe Villers Experimental Media Facility—this 3,600 square foot, three-story space is used for large-scale experiments and public demonstrations;
- Bartos Theatre—the 200-seat auditorium with sound system and film projection facilities is used for symposia, lectures, workshops, and classes;
- sound studios—for speech, acoustical and sound representation research;
- mechanical and electronic fabrication facilities;
- digital/analog multimedia development suites;
- holography lab—one of the largest collections of vibration-isolated precision optical table space in the country, totaling 260 square feet;
- fully-equipped digital/analog camera and range sensor studio.

Hardware

- IBM Power Visualization System—a 32 processor, high bandwidth array processing system for image processing, compression distribution;
- in-house developed "Cheops" data-flow processor for real-time audio, video, and holographic picture processing. This system also incorporates special-purpose processing modules built to analyze, correlate, and perform other processing functions on short-term projects;
- over 250 (more than one per researcher) workstations—DECstations, Alphas, HP9000s, SGIs, Sun Sparcstations, and IBM RS/6000;
- over 265 Macintosh and IBM PCs for programming, administration, and remote use;
- a terabyte of storage space distributed throughout the building.

Software

Image and sound processing libraries developed in-house include Matlab, Mathematica, Alias, and MPEG and JPEG compression routines. We are participants in the MPEG standardization group, and maintain reference coders and test model coders.

Content

Associated Press, Knight-Ridder/Tribune, and Reuters provide us with electronic news and picture sources. These news feeds are provided in support of our electronic news delivery and analysis research projects.

Fiber Optic Network Testbed

As 1996 began, the Media Laboratory was bringing to completion a three year $3.5M program to transform the Wiesner Building into a "fiber-to-everywhere" facility. This allows us to experiment with the cohabitation of signals in point-to-point and point-to-multipoint channels ranging from 1200 baud to 100 gigabits per second without making distinctions among audio, video, and data.

AMP Inc. provides cable, connectors, patch panels, and concentrator electronics, allowing the Lab to run Ethernet, FDDI, ATM, and future network technologies over the same plant and enabling—but not forcing—migration to higher speeds as necessary. Bay Networks, Inc. provides the routers and Ethernet switches. They enable diverse local-area networking technologies to communicate, and provide strong wide-area networking capability. Connectware, a wholly-owned subsidiary of AMP, supplies signal concentration electronics for Ethernet applications.

The network testbed provides a framework for investigating the impact of high-speed network technology on digital media. It will equip Media Laboratory researchers with a very high performance, extremely flexible and configurable network infrastructure. The network design features are as follows:

- very high density vertical fiber backbone (1,400 vertical fibers);
- generous coverage of all lab and office space with fiber to the desk (1500+ fiber drops);
- cable plant topology which allows any site within the Media Laboratory to be interconnected;
- flexible network layer organization which allows the rapid implementation of new network segments.

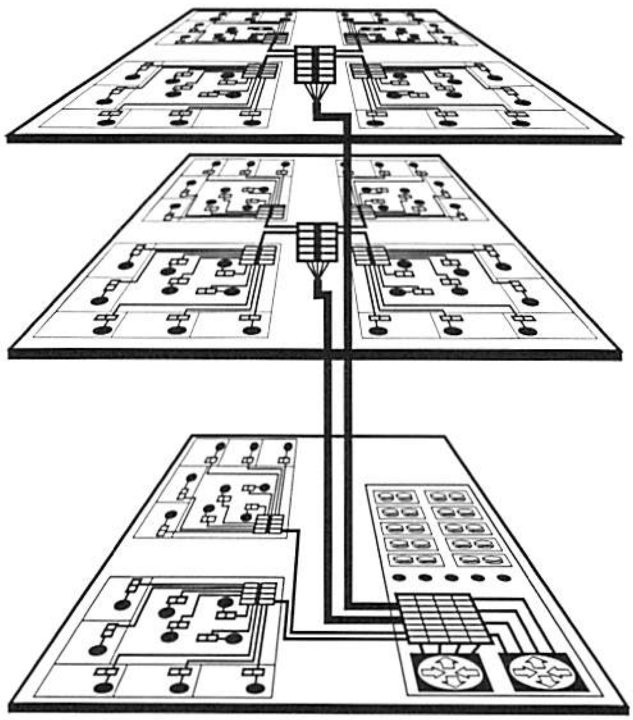
This network will allow the Media Lab to pursue high-bandwidth research, which would otherwise be infeasible.

The Media Lab has completed the following:

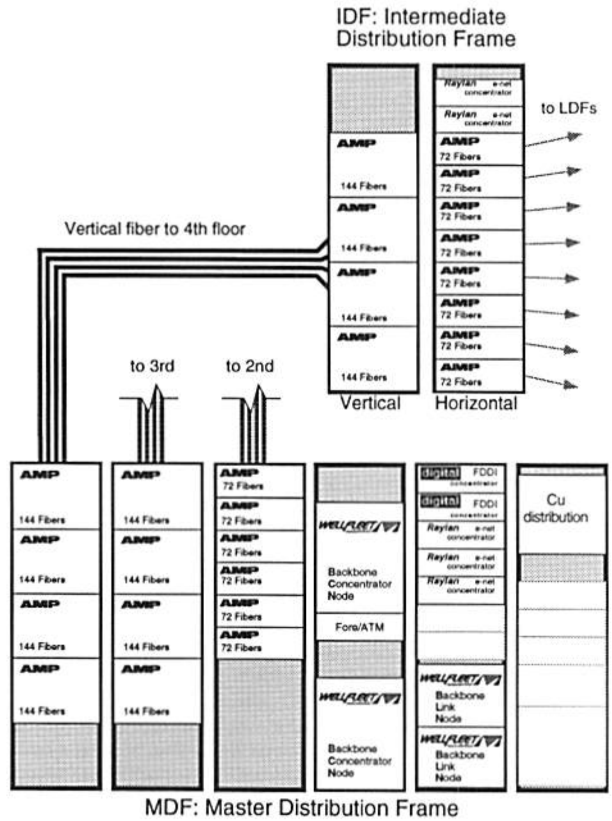
- the backbone phase of the fiber optic cable plant;
- most distribution frames (MDF, IDF, and LDF) including associated wall plates and portable outlets;
- connection of over 90% of the workstations and 60% of the Macintoshes to fiber;
- a network layer architecture design based on Bay Networks routers which allows for a smooth transition from a bridged to a routed environment. This architecture has been implemented and tested with Ethernet, FDDI, and ATM segments.

Schematic overviews of the Media Lab fiber optic network and of the network distribution systems are illustrated in the following Network Overview, Master Distribution, Intermediate Distribution, and Load Distribution diagrams.

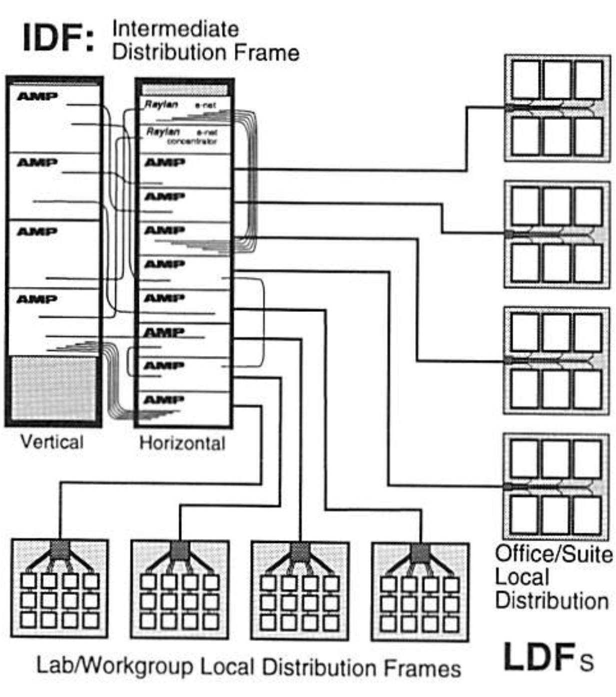
Fiber Optic Network Overview



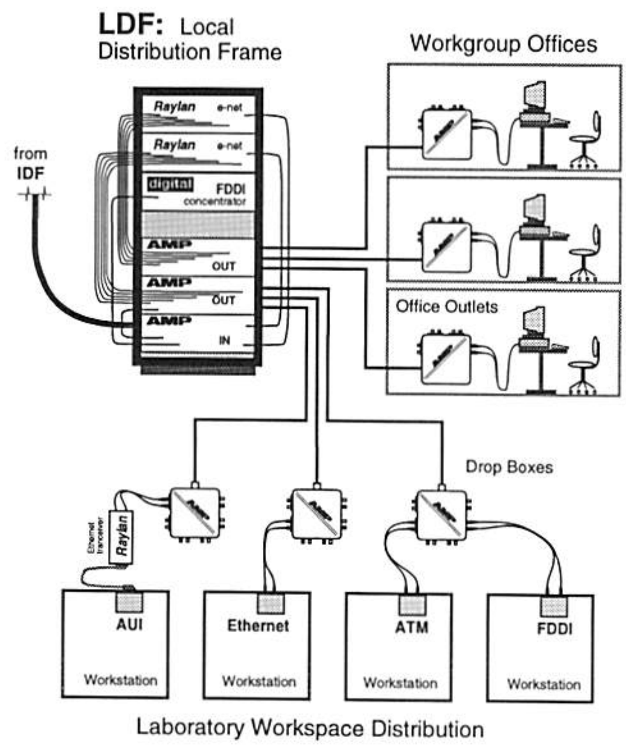
Master Distribution



Intermediate Distribution



Local Distribution



Massachusetts Institute of Technology

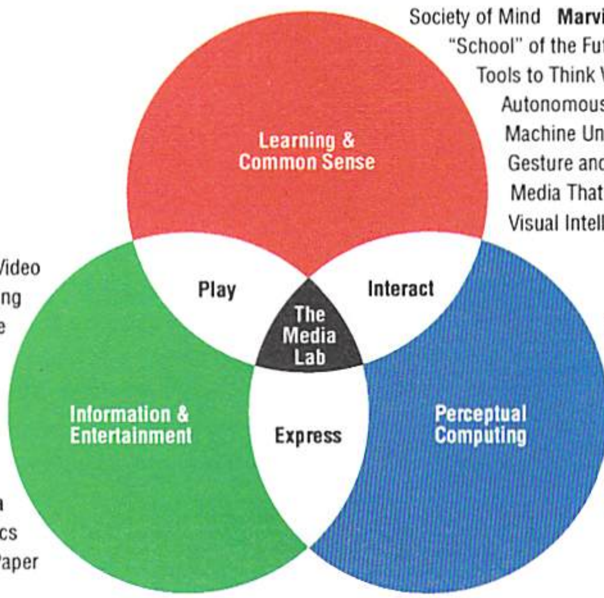
Media Laboratory

Projects

October 1995

Society of Mind **Marvin Minsky**
"School" of the Future **Seymour Papert**
Tools to Think With **Mitchel Resnick**
Autonomous Agents **Pattie Maes**
Machine Understanding **Kenneth Haase**
Gesture and Narrative Language **Justine Cassell**
Media That Learn **Henry Lieberman**
Visual Intelligence **Matthew Brand**

Andrew Lippman Structured Video
Stephen Benton Spatial Imaging
Walter Bender News in the Future
Tod Machover Opera of the Future
Michael Hawley Information Appliances
Neil Gershenfeld Physics & Media
Hiroshi Ishii Seamless Media Design
V. Michael Bove Object-Oriented Video
Glorianna Davenport Interactive Cinema
Ronald MacNeil Intelligent Graphics
Joseph Jacobson Electronic Paper



Cognition **Whitman Richards**
Looking at People **Alex Pentland**
Synthetic Listeners & Performers **Barry Vercoe**
Multimodal Interaction **Richard Bolt**
Affective Computing **Rosalind Picard**
Computers Understanding Action **Aaron Bobick**
Speech **Christopher Schmandt**

Communications and Sponsor Relations
20 Ames Street, Room E15-234
Cambridge, Massachusetts 02139
casr@media.mit.edu
http://www.media.mit.edu
Tel: 617/253-0338
Fax: 617/258-6264

Many of the MIT Media Laboratory research projects described in the following pages are conducted under the auspices of three sponsor-supported, interdisciplinary Media Laboratory consortia. These are:

News in the Future

The News in the Future (NiF) research consortium provides a forum for the MIT Media Laboratory and member companies to explore and exploit technologies that will affect the collection and dissemination of news. The goals include enhancing the efficiency of production, the timeliness of delivery, the convenience of presentation, and the relevance of editorial and advertising content to the consumer. NiF focuses on four areas: description of news by and for computers; observation and modeling of consumer behavior; presentation and interface design; and application. The consortium develops technologies for managing data, building linkages between news providers and consumers, and enabling new approaches to the look and feel of news content. It investigates the application of these technologies by means of experiments at MIT and field experiments set up in cooperation with individual member companies.

Television of Tomorrow

Television of Tomorrow (TVoT) explores the advent of digital television: one of the few realities in today's communications revolution. Tomorrow's television will be more than entertainment—it will be the *lingua franca* of commerce, education, research, and personal information. The research of this program is the foundation for distributed digital libraries, on-demand audiovisual entertainment, and personalized, content-aware networks. "Television" is used as a metaphor for its breadth and reach: the deeper issue is inventing new computational media that will have the capability to paint a canvas unique for each viewer. This consortium draws on an international group of sponsors who are information providers, channel-makers, and hardware manufacturers. Current research areas include: Media Bank, Structured Video, Structured Audio, Cheops, Holographic Video, Content Coding, Electronic Landscapes, and Story Representations.

Things That Think

This new consortium, officially innaugerated at the Laboratory's 10th birthday celebration, explores ways of moving computation beyond conventional sites, such as PCs or laptops, and adding intelligence to objects that are first and foremost something else. By sensing the movements or feelings of their owners—or by learning their owners' habits—common devices such as toasters, doorknobs, or shoes, will be able, in their own right and through communication with one another, to solve meaningful problems. By becoming truly responsive and unobtrusive, the information technology in the inanimate things around us will enhance the quality of daily living. Things That Think (TTT) brings together an unusual range of interdisciplinary talent and builds upon the close research partnership between the Laboratory and its sponsor community.

The information in this document is also available on the World Wide Web at http://www.media.mit.edu, where it will be updated periodically.

I. LEARNING & COMMON SENSE

1. Visual Intelligence

(Dr. Matthew Brand)
Every picture tells a story. How do we "see" that story? Our powers of vision depend on a rich body of knowledge that relates the appearance of scenes to their underlying causality. Seeing is largely a process of explaining sensations in terms of this knowledge. We aim to capture this knowledge, formalize it, and deploy it in intelligent computer-vision systems. The goal is to produce machines that are visually fluent and can cooperate with humans in activities such as design, editing, assembly, repair, and play.

2. An Eye for Design

(Dr. Matthew Brand)
As we build a semantics for vision, we are looking at the use of this visual knowledge in computer vision and in the generation of new and interesting visual experiences. The goal is to have computers design objects that are readable, informative, and visually appealing to humans. Projects include the automatic generation of graphics and layouts; graphical representations of the content of large text corpora; and a robot sculptor that will record its visual impressions by producing Calder-like mobiles.

3. Animated Conversation

(Professor Justine Cassell)
Because non-verbal signs are integral parts of the communicative process, we are designing a system that integrates gesture, intonation, and facial expression into multi-modal human-figure animation. In this project, appropriate speech, intonation, facial expression, and gesture are generated by rule from a semantic representation that originates in a goal-directed conversational planner. The output of the dialogue generation is used to drive a graphical animation of a conversation between two simulated autonomous agents. The two animated agents interact with one another, producing speech with appropriate intonation, hand gestures, and facial movements, such as head turns and nods. Future directions for this work include generating stories as well as conversation, adding a vision component so that the two agents can perceive each other's movements, and adapting the system to handle human-computer interaction.

4. Renga: the Cyberstory

(Professor Justine Cassell)
As part of our exploration of storytelling as a collaborative activity, we are developing a Web system that enables and encourages collaborative storytelling among children. The Web system is called "Renga," and it allows children around the world to collaboratively tell a story in real time. Renga (from the Japanese word meaning "linked poem" or "linked image") resembles the game played in primary schools where the class

sits in a circle, one child begins a story, and each of the other children, in turn, adds a sentence. Renga incorporates many of the linguistic, imaginative, and community aspects of the old-fashioned round-robin story, but also adds the capacity for children who are not in the same room—in particular for children in different countries—to share in the circle. The core of the Renga system is a Web page, although children without access to the World Wide Web may participate by e-mail and Minitel (in France).

5. Penelope, Story Weaver

(Professor Justine Cassell)
How can we take advantage of children's play styles to encourage them to try new kinds of toys and new kinds of technologies? In particular, how can we use the affinity of girls for storytelling and exploration of social relationships to draw them into new technology; and how can we use the affinity of boys for new technology to draw them into storytelling and exploration of social relationships? In the light of these questions, we are currently designing a toolkit that allows children to design their own computer games. The toolkit provides a template that allows the child to develop characters, a storyline, some narrative language, and a set of goals for a storytelling and problem-solving computer game. The child can then interact with the story-game that s/he has designed, playing the part of any one of the characters, or the part of an omniscient playwright.

6. Memory-Based Representation

(Professor Kenneth Haase)
Memory-based representation is a model of representation that uses analogies between descriptions to identify semantic roles and categories. Memory-based representations begin with broad and shallow representations which gradually become more structured and organized. They are an improvement over the conventional representations, which require substantial up-front design effort and adapt poorly to changes in representational assumptions. We have developed a number of very efficient algorithms for determining analogies between descriptions, and are working on applying these algorithms to large databases with millions of concrete descriptions. Most of our work has focused on automatically parsed databases of text, although now we are beginning to move into involving images and other sorts of descriptions.

7. Understanding News

(Professor Kenneth Haase)
Memory-based representation systems are being applied to comprehending, filtering, and summarizing news stories. News stories taken from wire services and other sources are run through a simple parser which annotates the text with phrase boundaries and possible relationships between phrases. This annotated text is then passed to the memory-based representation system and "understood" by identification of, and connection with, similar stories already in memory; preferences and queries are interpreted as partial stories which match incoming or recorded descriptions. Comparison of such

understood texts with texts previously read by a user allows user-specific summarization of new articles based on the real differences between articles. In addition to filtering incoming daily news, these tools provide an interface to large text databases and other sorts of databases (e.g., images and video segments) annotated with textual descriptions. One strategic advantage of this approach is that, in the worst case, it does as well as keyword matching (i.e., similar words indicate similar article). In the best case, it does as well as a human editor or selector.

8. Iconic Stream-Based Video Logging

(Professor Kenneth Haase)

Media Streams is an iconic logging system for video content which provides the descriptions used by storyteller systems, archival retrieval programs, content-based editors, and other systems that can take advantage of knowing the content of recorded video. The logger treats video as a stream with temporally bounded events, rather than as a set of clips with attached keywords; this allows the system to automatically "cut" the video to its own purposes. Video annotations are represented graphically to enhance data visualization and to enable logs to be shared among human and machine users. In addition, palettes of commonly used sets of iconic annotations streamline the logging of segments that are similar to segments seen before. The indexing of both the video itself (whose images are stored digitally), and of the icon palettes, connects to the facilities of a memory-based representation in the background.

9. Storyteller Systems

(Professor Kenneth Haase and Professor Glorianna Davenport)

Storyteller systems are sophisticated programs with deep and detailed knowledge of some particular domain or domains, with access to "media resources"—recorded video, sound, and text —regarding the domain. By combining these resources with synthesized graphical and textual representations, a storyteller system produces a story customized to what it knows—and what it learns—of a listener's background, preferences, and interests. These stories emerge dynamically as the system interacts with the user; questions and criticisms yield wholly new sequences of video, sound, and explanation in reply. Such systems transform the character of a publication: rather than producing epistles, one produces emissaries.

10. FRAMER: Knowledge Description and Sharing

(Professor Kenneth Haase)

FRAMER is a portable library for knowledge representation and inference being used in a variety of projects around the Lab. FRAMER provides a persistent object-oriented database with a simple inheritance mechanism and an embedded extension language (FRAXL) based on SCHEME. FRAMER data structures are easily shared between different hardware platforms (workstations, Macintoshes, PCs) and software platforms (C and LISP). Current work on FRAMER includes the development of a portable user interface API for FRAXL, a networked implementation supporting the distribution of programs and data, and integration of ongoing analogical representation work with FRAMER.

11. Beyond Precision and Recall

(Professor Kenneth Haase)
Current information retrieval systems are really only "document retrieval systems": once documents are retrieved in response to a query, the user still has to slog through them individually to extract the needed information. This model breaks down as the number of genuinely relevant documents grows. Even if a retrieval system has perfect precision and recall, queries for information may produce too many documents to search manually for needed information. Real information access systems need to organize the queries they return based on commonalities, essential differences, etc. Based on our text understanding and memory-based representation work, we are building information access systems which tell "stories,": combining, comparing, contrasting, and extracting retrieved documents.

12. Getting Computers to Know What They're Talking About

(Professor Kenneth Haase)
Human speech uses numerous cues to guide a listener's comprehension of the spoken word. Pause, intonation, inflection, and rhythm all combine to inform listeners of how the content of speech should be coordinated with memories and expectations. The absence of this information makes computer-synthesized speech notoriously difficult to understand. The computer doesn't know the connections between—or the context of—the words it is speaking, so the burden of this determination falls entirely on the human listener. We are exploring a distributed model for "online" speech production that combines semantic and contextual information to produce natural-sounding speech. We hope both to elucidate some of the processes underlying human speech production and to generate speech that is easier to understand, harder to misunderstand, and easier to listen to with partial attention.

13. Augmented Text

(Professor Kenneth Haase)
Augmented text systems combine the understanding of news with the understanding of the reader's context in order to produce ancillary information which places material into perspective. For instance, the augmented text for a description of a natural disaster in some distant locale would include annotations with respect to the reader's own community. Such annotations might include, "The affected area is about the size of Brookline," or "Sorghum is 75 percent of the country's agricultural output." Augmented text makes the news both more comprehensible and more immediate, enhancing its effectiveness and attractiveness.

14. Using Knowledge in Search and Retrieval

(Professor Kenneth Haase)
We are developing systems that help users find information in distributed networks like the Internet. Finding information in

response to a user's query involves two steps: information access, where the system finds information archives that might contain relevant information; and information retrieval, which is the process of searching such archives. Our aim is to develop systems that can match queries, not only to literally related, but also to conceptually related information. For example, if the user is interested in the topic of pollution, our systems would match this query to information archives about the environment. By making use of online dictionaries and semantic thesauri, we have developed automatic techniques to extract related concepts. These techniques have been incorporated into two programs, NetSerf and ImEngine. NetSerf is a resource discovery tool for finding relevant information archives on the Internet. ImEngine is used to construct semantic descriptions of pictures and video clips by processing their captions, and to retrieve relevant pictures and video clips using these descriptions.

15. Aesthetic Comprehension

(Professor Kenneth Haase)
Any full and effective model of human cognition must include aesthetic and artistic appreciation as a mode of human understanding. We are exploring models of aesthetic comprehension arising from our work on memory-based representations. Our chief thesis is that aesthetic experiences involve changes to the way our minds retrieve and compare experiences. Given this model and our current memory-based architecture, we can build machines whose exposure to particular works changes the way they index and match descriptions in the future. Our current prototype applies technologies developed for indexing and matching prose (news) texts and applies them to poetry. "Understanding" the poem involves drawing analogies based on textual and semantic cues and then using those analogies in the future.

16. Graphical Interfaces for Software Visualization and Debugging

(Dr. Henry Lieberman)
This project explores how modern graphical interface techniques and explicit support for the user's problem-solving activities can make more productive interfaces for debugging, which accounts for half the cost of software development. Animated representations of code, a reversible control structure, and instant connections between code and graphical output, are some of the techniques used.

17. Intelligent Technical Documentation

(Dr. Henry Lieberman)
Technical documentation for hardware and software is expensive to produce, often inaccurate, and inadequate. We are exploring a new approach to producing technical documentation in which an expert interacts with a simulation of a device, and the system automatically produces both English descriptions and visual illustrations.

18. Graphical Annotation

(Dr. Henry Lieberman)
People often communicate important knowledge by drawing and labeling diagrams. Why can't we also communicate

knowledge to a machine by using graphical indications of parts and structure, rather than by textual databases or programming languages? We are using computer-readable graphical annotation of images in a direct-manipulation editor to communicate relations that tell the system how to interpret and generalize user actions. We are also exploring voice input so that the user can explain actions to the machine as they are being performed.

19. Navigating in Very Large Display Spaces

(Dr. Henry Lieberman)
How would you browse a VERY large display space, such as a street map of the entire world? The traditional solution is zoom and pan, but these operations have drawbacks that have gone unchallenged for decades. Shifting attention loses the wider context, leading to that "lost in hyperspace" feeling. We are exploring alternative solutions, such as a new technique that allows zooming and panning in multiple translucent layers.

20. An Advisory Agent for Web Browsing

(Dr. Henry Lieberman)
We are building a new kind of agent that acts as a user's assistant in browsing the World Wide Web. Many current Web tools perform searches for the user, but our approach is to consider search for information as a cooperative venture between the human user and an intelligent software agent. Rather than search a pre-indexed portion of the Web according to user-stated keywords, the agent, Letizia, infers interest implicitly from observing user actions and tries to stay just a few steps ahead of the user, searching the user's immediately accessible links dynamically.

21. Instructible Agents

(Dr. Henry Lieberman and Dr. David Maulsby)
Agent software can perform tasks automatically on behalf of a user, but how does the agent come to learn what the user wants? Sometimes the agent can learn just by observing user behavior, but there may also need to be interaction where the user instructs the agent more explicitly. The instructibility aspect is the focus of this project. The user may present examples of behavior that the agent should follow and give advice to the agent as to how the examples should be interpreted. The agent must give feedback to the user so that the user understands what the agent knows and is capable of doing. Multimodal interaction is important in both the instruction and feedback.

22. Agent-Application Communication

(Dr. Henry Lieberman and Dr. David Maulsby)
Current experiments in agent software rely mostly on domain-specific applications that have been programmed from scratch, or explicitly modified with an agent in mind. Is it possible to make a toolkit or protocol that would allow an agent to communicate and control applications that have been constructed more conventionally? Can the agent "take the place" of the user in the interface? Can the agent have access to the application's data and behavior? Will commercial "inter-application communication" mechanisms suffice? What is the

division of labor between the agent and the application? This work will explore these questions.

23. Graphics by Example

(Dr. Henry Lieberman and Dr. David Maulsby)
Experts in visual domains such as graphic design are fluent in the generation and critique of visual examples. We are combining representation and learning techniques from artificial intelligence with interactive graphical editors, to create a "programming by example" system that can assist designers in automating graphical procedures.

24. Breaking the "Berlin Wall": Integrating AI and Graphics Programming

(Dr. Henry Lieberman and Dr. David Maulsby)
An obstacle to building systems that incorporate AI and interactive graphics is that much software in these domains is incompatible. AI often uses flexible, symbolic languages like Lisp, whereas efficiency-critical code may use C or C++. We are exploring ways of breaking down this barrier to create an environment that can combine software written from multiple perspectives and get the best of both worlds.

25. ACT (ALIVE Critter Training)

(Dr. David Maulsby and Dr. Henry Lieberman)
Agents in the ALIVE system (see project number 26) interact with users according to pre-defined, stimulus-response behaviors. To enrich users' experience of ALIVE, we will enable them to put their own creations into the virtual environment—in particular, to create new objects out of virtual LEGO components, and to teach ALIVE agents new behaviors. The ACT system adapts techniques from "programming by example" (see project number 23) so that agents can learn to imitate users' behavior or modify their responses to please the user. For instance, an ALIVE character would watch the user construct a LEGO building and help out by bringing pieces or even installing them. Since ALIVE agents know more than users about some aspects of the virtual environment, they can teach as well as learn.

26. ALIVE—Interacting with Animated Autonomous Agents

(Professor Pattie Maes)
To date, the cumbersome nature of the equipment and the limited nature of the interaction has limited the range of applications of virtual environments. ALIVE (Artificial Life Interactive Video Environment) is a novel system that allows wireless, full-body interaction between a human and a rich graphical world inhabited by autonomous agents. The ALIVE system provides more complex and very different experiences than traditional virtual reality systems. In particular, we are exploring novel applications in the area of training and teaching, entertainment, and last but not least, digital assistants or interface agents. Current ALIVE worlds that the user can experience include a virtual dog with whom the user can play and video-game creatures with whom the user can interact. A synthetic animated aerobics teacher, who gives the user detailed personal feedback, is currently under construction.

27. Modeling Autonomous Agents

(Professor Pattie Maes)
An autonomous agent is a computational system that inhabits a complex, dynamic environment. The agent can sense, and act on, its environment, and has a set of goals or motivations that it tries to achieve through these actions. Depending on the type of environment, an agent can take different forms. Autonomous robots are agents that inhabit the physical world; computer-animated characters inhabit simulated 3-D worlds; while software agents inhabit the world of computer networks. The same basic questions are studied in these different domains: How does an agent decide what to do so as to progress toward its goal (the problem of action selection); how does an agent learn from experience; and how can multiple agents collaborate? Our research team develops new techniques and algorithms to address one or more of these issues.

28. Modeling Synthetic Characters for Interactive Storytelling

(Professor Pattie Maes)
This project applies techniques from a range of disciplines (artificial intelligence, artificial life, animation, literature, theater) in developing autonomous, interactive, synthetic characters for virtual environments. In particular, we explore how to make such characters act and react in life-like ways, how to give them personality and emotions, and, in general, how to make them entertaining and engaging to a user. The resulting characters are used in interactive storytelling applications —story systems that dynamically adapt to a user's inputs. For example, we have used the ALIVE virtual environment to allow a user to enact one character in a computer-animated story in which the other characters are synthetic.

29. Software Agents

(Professor Pattie Maes)
This project applies artificial intelligence techniques to the field of human-computer interaction. In particular, techniques and systems developed in the areas of autonomous agents and common sense representation are combined to implement "software agents": interfaces that provide assistance to a person engaged in the use of a particular computer application. Interface agents differ from current-day interfaces in that they are more autonomous (performing many of the time-consuming, more mundane tasks the user normally would have to perform), more intelligent (learning from the user by observation and querying), and more personalized (customizing according to the user's goals, needs, preferences, habits, and history of interaction with the system). The project focuses on generic techniques for building interface agents, and has produced agents for a wide variety of applications including electronic mail and scheduling software.

30. Agents that Reduce Information Overload

(Professor Pattie Maes)
This project attempts to deal with the problem of information overload. We are building "software agents" that make personalized suggestions to a user for items the user may want to select (news articles, videos, music, television shows, classified advertisements, etc.). The project employs two

different techniques: content-based filtering and collaborative filtering. The former technique is used to detect patterns among the items liked or disliked, based on keywords and other features of the items. The second technique is used to detect patterns among different users and to make recommendations to people, based on others who have shown similar tastes. It essentially automates the process of "word of mouth" to produce an advanced, personalized marketing scheme. We have built filtering agents for applications such as news filtering, music recommendations, book recommendations, and recommendations for World Wide Web documents. We are currently building a system for "intelligent classified ads."

31. Agents that Help Form Groups of People with Similar Interests

(Professor Pattie Maes)

This project is developing an agent that finds people who have never met, but share similar interests, and introduces them to each other. Such introductions can automatically form interest groups and coalitions, and can be used to locate someone knowledgeable in a particular area. Each participating user runs a copy of the agent, and these individual copies find each other as appropriate on the network and begin the introduction process. The project is an experiment in creating a decentralized, fault-tolerant application that handles potentially sensitive information (such as people's mail, their personal files, or lists of their particular interests) in a responsible and privacy-protecting fashion, using cryptographic and other techniques. The eventual goal is ubiquitous deployment across the Internet.

32. Society of Mind

(Professor Marvin Minsky)

Professor Minsky continues to develop the theory of human thinking and learning called the "Society of Mind," which tries to explain how various phenomena of mind emerge from the interactions among many different kinds of highly evolved brain mechanisms. In this way we can account for many aspects of common sense, imagination, and reasoning by analogy, as resulting from negotiations among systems that use different ways of representing knowledge. Similarly, it appears that we can explain many of the regularities found in natural languages as consequences of how those representations work—rather than as constraints which are externally imposed on interpersonal communications. This approach also suggests that some of what we call "emotions" are mechanisms required for managing conflicts among competing goals. We may need to construct similar systems when we begin to build smarter and more versatile machines.

33. Project Headlight

(Professor Seymour Papert and Michele Evard)

In 1984, we began a partnership with the Hennigan School, a multicultural, public elementary school in Boston. In collaboration with the school and with IBM, we helped install what was then an unusually technologically rich learning environment, with enough computers to allow each student approximately an hour a day of access. Many early ideas about the use of computers in the context of an essentially traditional school were forged and tested in this setting. Recently, most activities are individually designed and implemented by teachers. Michele Evard works with several of them on projects involving telecommunications, student consulting networks, and news.

34. Learning in Multicultural Settings

(Professor Seymour Papert and Paula Hooper)
For several years, we have focused on issues related to gender, race, culture, and cognitive styles. One setting for this research is Paige Academy, a small, independent Afrocentric school in the Roxbury section of Boston. This setting provides an organizationally and culturally different context for developing new ideas about learning.

35. Mathematical Thinking

(Professor Seymour Papert)
We are studying the nature of mathematical thinking, and examining how the availability of new computational tools can support new types of mathematical investigations and new approaches to mathematics education.

36. Telling Why Tops Don't Topple

(Professor Seymour Papert)
The 1995 slice of a long-term project aimed at developing the "Art of Discussing Complex Ideas in New Media" focuses on how to carry on a discussion with someone far away on topics like tops and gyroscopes. The natural, traditional way to convey one's intuitive ideas on such subjects would involve pointing to the way the top wobbles, feeling what it is like to spin the top or hold the gyro, or drawing a picture. The goal is to understand how people explain and communicate ideas about complex phenomena, both through face-to-face, real-time communication, as well as by using electronic communication tools such as e-mail, teleconferencing, and other media.

37. Escher's World

(Professor Seymour Papert, Professor Whitman Richards, and David Williamson Shaffer)
Escher's World is a place where students create art and mathematics simultaneously in a studio setting. Making mathematics in such an expressive environment questions the very nature of what we mean when we say something is "mathematical." When children use mathematics as a tool for self-expression, they discover the visual, intuitive, and open-ended aspects of mathematical inquiry that are often missing from traditional mathematics classrooms. In this way, Escher's World explores how media technology dissolves the boundaries between traditional school "subjects" —and how these changes, in turn, force us to re-examine our understanding of what it means to think and to learn.

38. Constructionism

(Professor Seymour Papert and Professor Mitchel Resnick)
We are developing "constructionism" as a theory of learning and education. Constructionism is based on two different senses of "construction." It is grounded in the idea that people

learn by actively constructing new knowledge, rather than having information "poured" into their heads. Moreover, constructionism asserts that people learn with particular effectiveness when they are engaged in constructing personally meaningful artifacts (such as computer programs, animations, or robots).

39. Technological Fluency

(Professor Seymour Papert and Professor Mitchel Resnick)
We are creating a set of activities to support the study and development of technological fluency in pre-college students. Just as fluency in language means much more than knowing facts about the language, technological fluency involves not only knowing how to use new technological tools, but also knowing how to make things of significance with those tools, and (most important) developing new ways of thinking based on use of those tools.

40. Programmable Bricks

(Professor Mitchel Resnick, Professor Seymour Papert, and Fred Martin)
We are extending the child's construction kit, building computational power directly into LEGO bricks. With these Programmable Bricks, children can spread computation through their worlds, creating their own "Things That Think." Children are using these new bricks (and related sensors) to build autonomous robots; to create "active rooms" (for example, making the lights turn on whenever anyone enters the room); and to collect data from their everyday activities (for example, counting the number of steps they take in a day). We believe these activities will provide children with a new image of (and a new sense of control over) computation.

41. New Programming Paradigms

(Professor Mitchel Resnick and Professor Seymour Papert)
We are developing new "paradigms" for computer programming, hoping to change the ways that people (particularly children) use and think about computation. For example, we are adding multiprocessing and graphical-programming capabilities to the Logo programming language. These new paradigms notably extend the types of projects that children can work on (e.g., making it much easier for children to create their own video games); they also help children develop new ways of thinking about certain mathematical and scientific concepts.

42. Neighborhood Networks

(Professor Seymour Papert, Professor Mitchel Resnick, and Alan Shaw)
Many computer-network projects aim to support "virtual communities" of users who are geographically separated. By contrast, this project aims to support, strengthen, and sustain relationships among people who live near one another in urban communities. We are developing new network software that supports real-time interactions among users of a collaborative database, creating a new type of "meeting place" for members of the community. As examples, in an economically

disadvantaged section of Boston, neighbors are using this system to organize a food cooperative and a crime watch, and to read online articles from *The Boston Globe.*

43. Decentralized Thinking

(Professor Mitchel Resnick)
The world is full of decentralized systems (such as bird flocks, market economies, and the Internet), but most people have difficulty understanding the workings of such systems, often assuming centralized control where none exists. We are developing a new modeling environment, called StarLogo, to help people (particularly pre-college students) explore the workings of decentralized systems. With StarLogo, for example, a user can write rules for thousands of "artificial ants," then observe the colony-level behaviors that arise from all the interactions.

44. Toys That Think

(Professor Mitchel Resnick, Fred Martin, and Rick Borovoy)
We are developing a new generation of intelligent toys with computational and communications capabilities embedded inside. For example, in the Dr. LEGOHead project, children can snap together different computationally active body parts of a LEGO character. Because Dr. LEGOHead "knows" how he has been configured, he can react accordingly. By combining computational behaviors and recognizable forms into high-level building blocks, the project makes it easy and fun for children to start programming creatures who can respond to their environment. In the process, children can explore what it means to think in a social and technological context.

45. Learning in Virtual Communities

(Professor Mitchel Resnick and Amy Bruckman)
We are studying how online communities might change the way people (particularly children) learn, play, and think about themselves. In one type of online community, known as MUDs, participants not only "talk" and exchange messages, but they collaboratively construct the virtual world in which they interact. We have organized one such world (called MediaMOO) to enhance community among media researchers. We are starting a similar project (called MOOSE Crossing) for children. We believe such worlds could help children to become meaningfully engaged in reading, writing, and programming.

46. The Computer Clubhouse

(Professor Mitchel Resnick)
The Computer Clubhouse is an after-school learning center, where young people (ages 10 - 16) use computers to work on extended projects related to their own interests and experiences. The Clubhouse focuses on youth from under-served communities, who would not otherwise have access to technological tools and activities. At the Clubhouse, participants design their own computer graphics, robots, video games, interactive newsletters, music, simulations, multimedia presentations, and animations. We are studying how new "learning communities" can emerge in such settings.

47. Children and News

(Professor Seymour Papert, Professor Mitchel Resnick, Walter Bender, and Michele Evard)
This research explores new ways for children to use news as a context for interaction and learning. We are studying children's views of traditional forms of news, and creating new news-related tools and activities for children. In particular, we are interested in new ways for children to "produce" and discuss news, not just "consume" it. With one of our systems, children can create their own online newsgroups, discussing ideas with other children who share similar interests.

II. PERCEPTUAL COMPUTING

48. Dynamic Scene Annotation

(Professor Aaron Bobick)
In a dynamic scene, what is in the image is less important than what is happening in the scene. We are developing dynamic description mechanisms capable of extracting the important aspects of the behavior or motion present in a scene. Two domains we are exploring are charting football plays and extracting choreography from a ballet sequence.

49. Smart Cameras

(Professor Aaron Bobick)
We are developing techniques that will allow a camera control system to understand enough about a dynamic scene so that it will be able to respond to requests for specific camera shots. Such a system must understand enough about the intent of a director, the semantics and dynamics of a scene, and the computer vision techniques capable of localizing specific objects and action, to decide where to aim a camera. The integration of this knowledge requires a common framework for representing the action in the scene.

50. Multimodal Natural Dialog

(Dr. Richard Bolt)
People in each other's presence communicate via speech, gesture, and gaze. The aim of this research is to make it possible for people to communicate with computers in essentially the same way. This research explores speech, free-hand manual gesture, and gaze as combined input modes to the computer. One side of this effort is adapting technologies to capture inputs from the user: a speech recognizer, gesture-sensing gloves, and a head-mounted eye-tracking system. These technologies are off-the-shelf, and as more efficient, less obtrusive technologies emerge, they will be assimilated into the work. The other side of the effort involves the creation and elaboration of the software intelligence to interpret input from speech, hands, and eyes, and to map to an appropriate response in graphics and speech or non-speech sound.

The main expected outcome from this research is that computer-naive people (read: most of the world) will be able to use everyday social and linguistic skills to access computers and computer-based media.

51. Looking at People

(Professor Alex Pentland and Professor Aaron Bobick)
This large, multi-year research project is composed of several different sub-projects including: real-time tracking of people's body positions as they point and move about in the work environment; gesture and expression recognition; and continued development of our real-time face recognition system based on "eigenfaces." Currently there are two "test bed" applications of this technology: a real-time virtual reality system called ALIVE (with Professor Pattie Maes) and a "smart" teleconferencing system.

52. Wearable Computing

(Professor Alex Pentland, Professor Michael Hawley, Professor Rosalind W. Picard, and Thad Starner)
To date, personal computers have not lived up to their name. Most machines sit on a desk and interact with their owners for only a small fraction of the day. Smaller and faster notebook computers have made mobility less of an issue, but the same, staid user paradigm persists. Wearable computing hopes to shatter this myth of how a computer should be used. A person's computer should be worn, much as eyeglasses or clothing are worn, and should interact with the user based on the context of the situation. With heads-up displays, unobtrusive input devices, personal wireless local area networks (BodyNet), and a host of other sensing and communication tools, the wearable computer can act as an intelligent assistant, whether it be through a Remembrance Agent, augmented reality, or intellectual collectives.

53. Video and Image Libraries: Representation and Retrieval

(Professor Rosalind W. Picard and Professor Alex Pentland)
One of the most significant problems with multimedia technology is that you can't find what you want. This is because, unlike text-only systems, you can't ask a computer about the contents of images or video. For instance, you can't ask the computer to "find another video clip like this one, but shot from another angle," or "find a video clip of me on the beach." We are working to solve these problems by making computers able to "see" the contents of images and video.

54. Affective Computing

(Professor Rosalind W. Picard)
Recent neurological evidence indicates that emotions are not a luxury; they are essential for "reason" to function normally, even in rational decision-making. Furthermore, emotional expression is a natural and significant part of human interaction. Whether it is used to indicate like/dislike or interest/disinterest, emotion plays a key role in multimedia information retrieval, user-preference modeling, and human-computer interaction. Affective computing is a new area of research focusing on computing that relates to, arises from, or deliberately influences emotions. The focus of the present project is on giving computers the ability to recognize affect.

Current applications include better learning systems (computer recognizes interest, frustration, or pleasure of pupil), and smarter "things" such as a steering wheel/seatbelt that sense when a driver is angry or incapacitated.

55. Virtual Bellows for Video

(Professor Rosalind W. Picard)
Artists exercise viewpoint freedom with cubism and collage, and photographers flex their camera bellows; both desire to express what they see from multiple perspectives. This research harnesses the power of perspective mathematically, allowing one to extract and modify perspective in video. Applications include image mosaicing, high-resolution digital cameras, high-resolution printing, and recognition of video scene changes and camera motion.

56. Perceptual Similarity Measures

(Professor Rosalind W. Picard)
People are great at identifying similar patterns in pictures, sound, or human behavior. But how they do it remains a mystery. Based on results from tests of how humans recognize visual patterns, we are building computer models to mimic recognition of perceptual similarity. Particular attention is given to how humans interpret directionality, periodicity, randomness, contrast, translation, rotation, perspective, and scale in natural scenes.

57. WearCam: Video Orbits for Visual Memory

(Professor Rosalind W. Picard)
We are building a wearable, wireless, head-mounted video camera (WearCam) for diverse uses such as a hands-free sports camera or a system for augmenting visual memory. However, the digital video processing problems are immense. This research explores combinations of physiological signals from "Affective Computing" with signal processing algorithms from "Video Orbits" for reducing the amount of video that must be processed, and assisting the user in deciding what video should be "remembered."

58. Semantic Image Modeling

(Professor Rosalind W. Picard)
If I state, "Atlanta is in Cincinnati today," it is unlikely you will think I am coherent. If, however, we are talking baseball, then the sentence is very clear. The context makes the interpretation not only easier, but possible. Similarly, with pictures, if you see blue at the top then it's probably sky. The goal of this work is to begin setting up two-way interaction between available contextual information and the models used to represent visual information. The ultimate goal is the one Shannon missed: putting semantic meaning into "information" theory.

59. Computers and Telephony

(Christopher Schmandt)
Computer workstations can provide a much-needed user interface to advanced telephony functions, provided a path exists between the workstation and switch. Controlling call set-up from a user's workstation allows a greater degree of personalization and dynamic call-handling, both for outgoing

and incoming calls. This project is being implemented in the ISDN environment of MIT's campus telephone network, using Phoneserver, a computer network interface to Basic Rate ISDN switching.

60. Desktop Audio

(Christopher Schmandt)
This project explores software architectures and user interfaces to voice as a computer data-type, as well as a command channel. Its goal is to make speech ubiquitous to a range of applications (for instance, editing a telephone message to include annotation of a text document). Related issues include object-oriented manipulation of multiple media "selection" (or "clipboard") data between processes, and a client-server architecture allowing multiple applications to share audio resources.

61. Voice Interfaces to Hand-Held Computers

(Christopher Schmandt)
This project considers the role of speech interfaces to very small, portable computers which manage voice as an inherent data type. One component is a hand-held audio note-taker, with a speech-recognition interface. Another is a portable "news radio" designed for occasional connection to a digital audio network to load new stories, and launch queries for stories about which the listener would like to hear more.

62. Telephone-Based Voice Services

(Christopher Schmandt)
This project explores the utility of voice in a range of applications offering services to users of the telephone network. Topics being examined include voice mail, speech synthesis of electronic mail, access to calendars and Rolodexes, and speech-based user interface to call-processing features, such as variable call forwarding. Visual- (on the workstation) and speech- (over the telephone) based applications offer differing views of the same underlying databases in an office environment.

63. Conversational Interfaces

(Christopher Schmandt)
This work uses combined speech input and output to converse with a user seeking to control the computer, or to access information from it. Conversational techniques allow the computer gracefully to limit the vocabulary that is likely to be spoken, facilitating speech recognition. Discourse techniques aid error detection and correction. Domain knowledge and learning about user preferences allow for determining a user's needs more efficiently. Listening for various forms of back-channel response from the listener permits a talking computer to gauge user interest better, and determine at what level of detail to describe the requested information.

64. Interactive Radio

(Christopher Schmandt)
Radio news programs are most valuable when hands and eyes are busy, such as during the morning commute or while working in the kitchen. But radio news is presented in small segments and only at scheduled times. This project seeks to

compile audio news from a number of sources, segment stories based on acoustic and possibly semantic cues, and present these on demand in an interactive environment, such as over the telephone or on a computer workstation.

65. Voice As Data

(Christopher Schmandt)
In order to utilize digitized speech (a telephone message, recording of a meeting or lecture, etc.) more effectively as a computer-data-type, it is important to be able to understand speech structure and allow rapid scanning of the audio data. Structure is derived from pauses, speaker changes and turn-taking, emphasis, and intonational cues. Presentation may be serial, including audio time-compression, or parallel, using both ears to listen to multiple audio streams simultaneously.

66. Acoustical Cues to Discourse Structure

(Christopher Schmandt)
This work seeks to apply acoustical analysis to a hierarchical model of discourse structure so as to allow for the division of an audio recording into layers of detail. This "outline" structure is then used in interactive presentations, which allow a listener to browse at various levels of detail.

67. Spatial Listening for Auditory Presentation

(Christopher Schmandt)
This project uses spatialized audio and presentation of audio recordings from different points in space around the listener's head to enable more efficient listening. Examples include:

• browsing multiple streams of audio simultaneously, with computer-enhanced, selective-attention modeling and automatic notification of salient events on secondary channels;

• spatial representation of a single audio recording to map time into position for more effective dynamic browsing; and

• auditory "information landscapes" which allow a listener to move in a virtual space among recordings.

68. Synthetic Performers

(Professor Barry Vercoe)
We have shown that computers can exhibit real-time musical behavior similar to that of skilled human performers. Our live violinist accompanied by a computer-driven piano has been widely viewed on public TV. This research uses real-time digital audio processors to explore the music-cognitive issues that arise when a computer is put in the position of real-time, highly sensitive human interaction.

69. Synthetic Listeners

(Professor Barry Vercoe)
How can machines hear what humans hear? By modeling the human auditory processes such as sound localization, multi-source signal separation, and pitch tracking, machines can exhibit elements of sound recognition ranging from audio feature detection and texture recognition to generalized auditory scene analysis. This enables machines to know who is in a room, where they are, and what they are probably doing.

70. Synthetic Spaces

(Professor Barry Vercoe)
Although simulation of 3-D audio space using just two loudspeakers has recently been enhanced by methods of transaural audio, the technique is fragile and sensitive to listener position. We are developing a generalized solution that would not impose such restriction. Dynamic Generalized Transaural Audio will employ a small number of speakers to enable multiple listeners to simultaneously receive the cues of arbitrary acoustic spaces as they move freely about the room.

71. Cognitive Audio Processing

(Professor Barry Vercoe)
Using real-time software sound synthesis and analysis, we are investigating how humans perceive and quantify music and audio information in cultural contexts. This involves computer-assisted identification of source type, intonation, rhythmic and tonal structure, and emotional content, within Western and non-Western traditions.

72. Structured Audio Transmission

(Professor Barry Vercoe)
Moving sound over the Internet is painful for a network that is 100 times slower than real-time quality audio. We are developing a sound-object encoding scheme (Netsound, a kind of PostScript Audio) that enables sound descriptors and schedulers to be sent efficiently over the Internet, then reconstructed in real-time at the receiver site.

III. INFORMATION & ENTERTAINMENT

73. Salient Stills

(Walter Bender)
A salient still is a photograph created from a video sequence. It contains both the context and the detailed content from the sequence. The process creates a data representation that consists of video pans, tilts, and zooms warped into a continuous space/time volume. A high-resolution, panoramic still image is extracted from this representation. The still image has both the wide view field captured by the short focal-length frames and the details captured by the long focal-length frames.

74. Color Semantics

(Walter Bender)
We are exploring the role of color alignment in the preservation of the experience of color. Central to this investigation is the formulation of a theory of color alignment and its measurement. Objective quantification of color-relatedness is desirable since it allows precise specification of color in relation to its surrounding visual context and state of visual adaptation. A secondary theme of this research is the role color alignment plays in the generation of expressive energy in color combinations. Expressive energy in color combinations can be

predicted based on the selection of color alignments. We are applying this work to the measure of degree-of-alignment between the window and background in a workstation. This work will provide guidelines for effective selection of window, font, and background colors for any given application.

75. Doppelgänger: Knowing the Individual

(Walter Bender and Jon Orwant)
Just as a system should "know" the data, it should also be cognizant of the user. The more the system knows about the user, the better able it is to make sense of the ambiguities and inconsistencies inherent in human communication. Our work in user-modeling involves the full exploitation of the user's computational environment so that information normally provided by the computer (e.g., idle time, schedule information, electronic mail subscriptions) and other more esoteric information (e.g., physical location tracking systems, eye-tracking systems, speech manipulation, electronic newspapers, model-building cameras) can be integrated to construct dynamic, individual, user models that change over time as the user changes and the system learns more about the user.

76. Data Hiding

(Walter Bender)
Data hiding, or steganography, is the process of embedding data into images and audio signals. The process is constrained by the quantity of the embedded data, the need for invariance of the data under conditions where the "host" signal is subject to distortions (e.g. compression), and the degree to which the data needs to be immune to interception, modification, or removal. We are exploring both traditional and novel techniques for addressing the data hiding process, as well as evaluating these techniques in light of three applications: copyright-protecting, tamper-proofing, and augmentation-data embedding.

77. FishWrap

(Pascal Chesnais and Walter Bender)
FishWrap is an experimental, on-demand, personalized, self-organizing, electronic newspaper that features topical and geographic news, as well as on- and off-campus calendars available to MIT undergraduate students, Media Lab staff, and faculty. We are exploring how various emerging technologies affect the news industry. As part of the experiment, students create a customized news presentation drawing from a large pool of information that has been sorted by topical and geographical interests. The information is gathered from traditional news sources (*The Boston Globe*, AP, Reuters, Knight-Ridder) and the local MIT community (students, deans, News Office, clubs, academic departments). FishWrap is delivered to users on-demand using a World Wide Web browser (http://fishwrap-docs.www.media.mit.edu/docs/).

78. News Games

(Walter Bender, Vadim Gerasimov, and Jon Orwant)
News should be fun, relevant, educational, and provocative. We are investigating news "games" as an alternative vehicle for accessing news content. Our goal is to entertain while engaging the player in the editorial and advertising content.

Our News Games include:

- **Autocross**: a crossword-puzzle generator that automatically creates puzzles based on news articles.
- **News Totalizer**: a future-prediction game available on the World Wide Web that encourages players to make predictions on the weather, politics, and sports.
- **NIF-Fish**: a version of the popular card game. Each card contains a story or a picture from different news categories.
- **MAMMON**: a network service that provides a player with a virtual portfolio of stocks, currencies, and mutual funds to play the stock market.
- **Paragraphs**: a news-reader puzzle that challenges the player to arrange paragraphs from a news story in the right order.

79. Scaled-Up Holographic Video

(Professor Stephen Benton, Carlton Sparrell, and Ravikanth Pappu)
The world's first electronic holographic video display has established the principles of information reduction and image scanning, but scaling up to practical display sizes has posed significant electronic and electro-optical challenges. The parallelization of the computation, storage, and display has been shown feasible for 3" x 5" images, laying the groundwork for further scale-ups of image size.

80. Holographic Laser Printer

(Michael Klug)
Full-color, wide-angle, and large-size computer-generated hard-copy holograms still take considerable time to create. A "holographic laser printer" allows simpler hard-copy holograms to be generated in minutes instead of hours, automatically and without wet processing. Research topics include recording materials and processing, optical design, image processing and LCD display, and optical techniques for image noise reduction.

81. Full-Parallax Synthetic Holograms

(Michael Klug)
Although the rapid generation of holographic images usually sacrifices image variations with up-to-down viewer motions, there are some situations where this information must be included. This substantially complicates the design of image rendering and recording technology, but leads to more useful images for some applications. This research studies one- and two-step techniques for producing images that present up-to-down as well as side-to-side image parallax.

82. Video-Based Holographic Portraiture

(Michael Klug and Michael Halle)
Holographic 3-D portraits can be produced in a few minutes from a rapid video-camera scan of a subject. The data are processed by a novel distortion-correction algorithm and sent to a holographic laser printer. The portrait is available for viewing soon thereafter. The project combines digital video technology with advanced image processing software.

83. Electronic Architectures for Interactive Holographic Video

(Carlton Sparrell)
The rapid decoding and display of large amounts of holographic video data require novel electronic architectures and devices that allow high-speed parallel processing. These take advantage of the underlying principles of diffraction-specific computation to simplify the circuitry and programming that are required.

84. Medical Image Holography

(Professor Stephen Benton and Michael Halle)
MRI and CAT-scan cameras gather 3-D data, but holography offers the only way of examining those images in fully 3-D form. This project explores new image-processing, editing, and rendering tools that are needed to make these complex 3-D images quickly and accurately interpretable by physicians.

85. Minimum-Pixel Holograms

(Ravikanth Pappu)
Holographic image communication inevitably requires the generation, transmission, and display of massive numbers of pixels. This number can be reduced, however, by several orders of magnitude by taking the characteristics of the human visual system into account. Further consideration of the optical system parameters and the properties of spatial light modulators leads to guidelines for the design of holograms with a minimum number of pixels needed for a given set of image properties.

86. Holographic Signal Compression

(Ravikanth Pappu)
Holographic video signals rarely use the entire bandwidth allocated for their transmission, reflecting the usual limitations on image content and complexity. Adaptive and non-uniform sampling methods allow such signals to be represented with roughly half the number of bits usually considered necessary for transmission.

87. Edge-Lit Holograms

(Professor Stephen Benton and Michael Klug)
Conventional holograms require illuminators to be mounted on walls or ceilings near the hologram; edge-lit holograms are a new type of white-light hologram that allows the light source to be included within the mount itself, assuring a compact and carefully aligned illumination. This project explores the fundamental diffraction and imaging properties of these holograms, with a view toward making their images deeper, brighter, and clearer.

88. Chromatically Corrected Holographic Displays

(Arno Klein)
Illumination of conventional reflection and edge-lit holograms with white light normally produces a spectral blur that degrades the image resolution. "Pre-filtering" of the illumination by holographic gratings and lenses can largely compensate for these chromatic aberrations to produce a sharp, multi-color

image that presents vertical as well as horizontal parallax. Research issues include the optimal design of diffractive optical elements, and their best matching to various holographic image types.

89. Three-Dimensional Haptic/Visual Interface Studies

(Wendy Plesniak)
Integration of a particularly agile new force-feedback 3-D pointing device with holographic video and other 3-D visual displays allows new styles of volume-graphical interaction to be explored.

90. Autostereoscopic Video Displays for Multiple Viewers

(Professor Stephen Benton and Paul Christie)
Novel optical projection systems, combined with viewer location tracking, can provide separate left- and right-eye views to small audiences. This project combines research in liquid-crystal display technology, specular illumination optics, and agile illumination systems.

91. Cheops: Data-Flow Television Receiver

(Professor V. Michael Bove, Jr.)
The Cheops Imaging System is a compact, modular platform for acquisition, real-time processing, and display of digital video sequences and model-based representations of moving scenes. It is intended as both a laboratory tool and a prototype hardware and software architecture for future programmable video decoders. Rather than using a large number of general-purpose processors and dividing up image-processing tasks spatially, Cheops abstracts out a set of basic, computationally intensive stream operations that may be performed in parallel, and embodies them in specialized hardware. A resource management daemon gives automatic software parallelization, hardware scalability, and real-time scheduling. Eight systems have been built and are in use at the Media Lab and at various sponsor sites.

92. Video Coding with Segmentation

(Professor V. Michael Bove, Jr.)
Most digital video-coding methods use a very simple approximation to scene motion that breaks up images into arrays of square tiles, and assigns a two-dimensional motion vector to each. We are developing video-coding methods that use motion, color, and texture to segment scenes into coherent regions and compute more accurate motions for the regions. The result should be a more compact representation, better scene understanding, and the ability to compute images for arbitrary instants in time.

93. Computing with Reconfigurable Hardware

(Professor V. Michael Bove, Jr.)
In order to address compactly and inexpensively the processing needs of our advanced video and audio representations, we have built small computing devices from SRAM-based programmable gate arrays. These devices give us both the power of specialized hardware and full programmability. We are now developing appropriate software techniques for managing these systems.

94. Model Extraction from Uncalibrated Camera Views

(Professor V. Michael Bove, Jr.)
Pictures of scenes containing rectilinear objects (such as buildings, room interiors, and furniture) often contain enough perspective information to enable recovery of the 3-D location and orientation of planar surfaces from a picture, even when the location and focal length of the camera are not known. Not only can we build a model from one view, but we also can add detail and extent by incorporating additional views of the scene. As part of the process, we characterize and correct lens distortions.

95. Production, Distribution, and Viewing of Structured Video Narratives

(Professor V. Michael Bove, Jr. and Professor Glorianna Davenport)
This project focuses on the use of object-based video and audio modeling techniques as a development and presentation platform for personalized video stories. The intention of the personalization may reflect requirements of scalable displays, as well as requirements of audience interaction. Most recently, we have produced a video loosely based on the short story "The Yellow Wallpaper," which addresses issues of story structure, scripting, production, and post-production, including the linking of sound and picture.

96. Elastic Media

(Professor Glorianna Davenport)
Digital media environments encourage us to design computationally driven stories which can be personalized. The design of these stories must invite viewer input. Both the story and the computational architecture must support the notion of reconfigurability. Prototype stories allow us to explore form in relationship to systems design.

- **Boston: ReNewed Vistas**: An evolving documentary which focuses on the redesign of Boston's Central Artery and its impact on neighborhood life.

- **Jerome B. Wiesner 1915-1994: A Random Walk through the 20th Century**: A hyper-portrait of former MIT President and co-founder of the Media Laboratory.

- **Crossing the Street**: A moment in time is witnessed by a diversity of characters. (Kevin Brooks)

- **Lurker**: This "Thinkie" invites the viewer to participate by thinking in a story-specific style. Multiple participants share e-mail and movie information in order to help the hackers solve a mystery.

- **Two Viewpoints**: This story loosely based on the short story "The Yellow Wallpaper," for Cheops, combines places, characters, and is shot on a blue screen, into a synthetic set. The story can convey the point of view of either one of the characters. (with Professor V. Michael Bove, Jr.)

97. Video Stories: Computational Partnerships

(Professor Glorianna Davenport)
Movie editing is extremely time-consuming, so time-consuming, in fact, that few home movies are ever edited. By integrating the moviemaker's knowledge of content and craft into software, we seek to create better partnerships between storytelling technology and the human storyteller. Work includes modules for control (Video Streamer), contextual annotation and browsing (ConArtist), and directorial annotation of scalable, structured video (SAR).

98. Stories with a Sense of Themselves

(Professor Glorianna Davenport)
In order to maintain both a story and an intentional display dynamic in an interactive, multi-threaded setting, story concepts as well as story content must be represented to the computer. Story Agents (Kevin Brooks) interpret story and assemble a coherently structured presentation. ConTour (Michael Murtaugh) enables coherent contextual presentation.

99. Newsroom of the Future

(Professor Glorianna Davenport)
As news producers increasingly mix background material into their online services, new story forms and techniques emerge. In this project we examine problems of generating and browsing evolving databases, transitions in multi-dimensional story spaces, and the idea of the journalist's notepad.

100. Storyteller Systems

(Professor Glorianna Davenport and Professor Kenneth Haase)
Storyteller systems are sophisticated programs with deep and detailed knowledge of some particular domain or domains and access to "media resources"—recorded video, sound, and text—regarding the domain. By combining these resources with synthesized graphical and textual representations, a storyteller system produces a story customized to what it knows—and what it learns—of a listener's background, preferences, and interests. These stories emerge dynamically as the system interacts with the user; questions and criticisms yield wholly new sequences of video, sound, and explanation in reply. Such systems transform the character of publication: rather than producing epistles, one produces emissaries.

101. Real-Time Modeling and Characterization

(Professor Neil Gershenfeld)
Digital systems that are used to describe realities, experiment with possibilities, or realize fantasies require efficient algorithms for the real-time modeling and characterization of complex systems. Because first-principle models of nontrivial systems quickly become computationally intractable, we are studying the use of state-space reconstruction techniques to develop compact models of the solution space of a system from measurements of the nonlinear input-output relationship. One area in which we are exploring this promise is in new musical instruments, both because of its significance for their evolution and because they provide an extremely demanding environment which requires the integration of multiple degrees of freedom of real-time I/O with high-speed computing.

102. Interface Transducers

(Professor Neil Gershenfeld)
Intelligent processors continue to be put in dumb boxes, yet the ability of computers to solve useful problems is increasingly limited by the ability of the user to interact with the machine. The interface of many of the most historically successful technologies, such as a Stradivarius violin or a Gutenberg bible, resides in the sophistication of the physical interaction with the device. We are developing the transducers needed for modern technology to capture this sophistication and naturalness. First, we have found that user actions can inexpensively and unobtrusively be measured with millimeter and millisecond resolution by detecting the tiny currents that arise from the interaction between people and varying electric fields. The limits of this technology are being explored, including 3-D imaging, and the use of more sophisticated measurement strategies for biometric characterization of the material in the field. This technology removes the wires from the user, but still requires wires to the sensor. A second project is to develop the use of active remote far-field interrogation of passive near-field structures to detect the ID, location, and state of tags. Finally, we are investigating the mechanisms and applications of smart materials for providing tactile feedback.

103. Information, Computation, and Physics

(Professor Neil Gershenfeld)
Information exists as high-level meaning (ideas), and as low-level physical representation (bits). Although these two levels of description are usually clearly separated, many of the most serious constraints and exciting possibilities in information technologies are at this interface. Physical insights can be applied to help understand computers. We are studying the use of techniques from the study of nonlinear time series for optimizing computer performance, and in the thermodynamic limit, and how this leads to fundamental bounds on energy and entropy budgets for practical tasks. Physical insights can also lead to new paradigms for computation. We are exploring the feasibility of classical and quantum computation in bulk materials through resonant spin interactions, and we have developed nonlinear dissipative generalizations of linear feedback shift registers, which overcome some of their important limitations.

104. Information Appliances

(Professor Michael Hawley)
Tools and appliances of all sorts, from wristwatches and notebooks to concert grand pianos and home entertainment systems, are sprouting digital components. To interoperate harmoniously, and to ease the personal interface to a global information system, appliances need to communicate with each other. This project studies the languages and systems required for an open and scalable architecture.

105. BodyNet

(Professor Michael Hawley, Professor Alex Pentland, Professor Rosalind W. Picard, and Thad Starner)
Your interface to the worldwide network should be as convenient as a wristwatch and as comfortable as an old

sweater. Like jewelry and clothes, it should be so comfortable and convenient that it is always available, always connected. The BodyNet is a very local, wireless network that will integrate all the information appliances on your person: cameras, watches, calculators, notebooks, datebooks, checkbooks, credit cards. Advanced displays, phones, radios, and other communication systems will be woven together to form an intimate, pervasive interface to the Net.

106. The Library Channel

(Professor Michael Hawley)
The future of public libraries is in the living room: the wealth of great libraries will, in time, be digitized and made universally accessible through home and personal appliances, via a plurality of wired and wireless means. The Library Channel project involves the Library of Congress, National Geographic, WGBH television, and other partners in the Television of Tomorrow program. Our initial emphasis is on prototype interfaces to media banks.

107. Personal Embroidery

(Professor Michael Hawley)
We have built a custom embroidery language and tools to drive computerized sewing peripherals. Think of this as "Media Fab." Think of an embroidery machine as a 254-thread-per-inch multicolor printer, which outputs to garments instead of photocopy paper. And think of thread as potentially being conductive wire, or optical fiber. We are playing not only with custom clothing design, but also with fanciful embroidered circuitry.

108. Electronic Paper Books

(Dr. Joseph M. Jacobson)
Books with printed pages are unique in that they embody the simultaneous, high-quality display of hundreds of pages of information. The representation of information on a large number of physical pages, which we may physically handle and write on, constitutes a highly preferred means of information interaction. The major drawback, of course, is the non-changeable aspect of the printed information.

We are currently embarked on a project to construct an "electronic book" composed of multiple electronically addressable displays formed on real paper pages. The project is titled *Scientiae Liberorum*. It falls within a larger project in which we seek to incorporate highly unexpected and highly useful attributes in everyday items by exploiting micro-mechanical systems fabrication.

109. Media Bank

(Dr. Andrew Lippman)
The Media Bank is a distributed web of processors that contain audio and video bits, snippets, and movies. The environment is a community of peers (a hookup of equally capable computers) instead of video servers and settop boxes. The system is designed to synthesize programs on demand and on-the-fly from annotated and cross-referenced elements scattered throughout the domain. Protocols consolidate synchronous,

real-time communication with emerging data-highway architectures. Stored data are richly indexed and continuously evolving, with analysis and history being added each time they are accessed. Research addresses methods for guaranteeing uninterrupted assembly of movies, dynamic transcoding, interactive visualizations, and indexing techniques. Within the next few months, the system will migrate to an internal ATM network.

- **Media Browser**: Media Bank data exist in multiple formats and diverse styles: from MPEG to scripts; from postage stamps to HDTV. There are movies, still pictures, clips, sound bites, and words. A Mosaic-based browser navigates through this information space and overlays a naming structure and organization. Automatic page-generation depicts the current Media Bank content and object location. This extension of the World Wide Web to migratory and evolving data is becoming the control path and authentication mechanism for the Media Bank.

- **MicroMedia**: The Media Bank is designed to capture evolving dissection and annotation of its content. A movie enters the system as a single catalogue entry and grows into lists of scene descriptions, actors, ratings, and cross-references. New versions are created by both human and automatic actions. Each viewing can add data to the bank, and each pictorial transformation makes a new version. Current research topics focus on ways to exploit the analysis already generated for compression, to provide navigational tools through movies, and to reveal their content.

- **Television Interfaces**: A television receiver is little more than a Media Bank client. Recent work shows how we can download the user interface and control structures along with the information itself. We can show the "Saturday Night Movie" on the Saturday Night television set that bundles unique viewing options with the program. Client work includes real-time extensions to existing Web browsers such as Hot Java and Mosaic.

- **On-Demand Formatting**: In two years no one will be able to sell video coders; their value may lie solely in the amount of content encoded within them. In anticipation of multiple representations and encodings of the same underlying information, we are constructing a set of linked, real-time transcoding filters, each of which can be called upon as a Media Bank service to be performed remotely or within the user's machine. These include scalable transformers, turning movies into books, and re-compressing.

- **Salient Movies**: We are producing a browsable, matte-like layout of an entire movie by a two-step process. First, we analyze every frame of the movie, and cluster similar ones into groups. Then each cluster is assembled into a "Salient Still" that encapsulates the action and camera motion. An array of these clusters helps expose the pacing and fabric of the movie in a book-like format.

- **Format Taxonomies:** Two graduate theses are in progress addressing extensions to motion analysis for resolution enhancement of moving elements, locating foreground actions, and building coding hierarchies that automatically use the best algorithm for each scene element. Applications include distribution of video data in scrambled formats that allow a viewer to catalogue interesting pieces of lectures and movies, and to explore simultaneous broadcasts.

110. Dimensionalization

(Dr. Andrew Lippman and Henry Holtzman)
Movie sequences are mapped into 3-D models of the static elements located in the scene. The sequences may be viewed from any vantage point; content can be realistically post-produced. New work on camera position estimation allows the model to be augmented by highlighting the changes from one image to the next.

111. Identity in Online Communities

(Judith Donath)
Online communities are growing rapidly in size and importance, but it is difficult to gain a sense of the other participants: most of the social cues we use in real life are missing. The goal of these projects is to create online environments that support more complex and subtle interactions.

- **Electric Postcards:** The Electric Postcard makes it possible to use the World Wide Web for interpersonal communication. As in the real world, the user chooses the postcard, writes a message, and sends it off to the recipient. These postcards, however, are electronic. No physical card is sent—the image and the message are kept online. Because the Electric Postcards are part of the World Wide Web, the message contents can be hypertext: images and sounds can be embedded in the text and there can be links to any other data in the Web. Available on the Web since January 1995, the Electric Postcard has proved to be quite popular—by the sixth month of use, 200,000 cards had been sent. This work is a first step in an investigation of the Web's complex interconnections as part of the vocabulary of electronic expression.
- **The Sociable Web:** The World Wide Web is a popular place, yet browsing through it is a solitary pursuit: one is unaware of the presence of fellow explorers. We are developing a server and client that allow Web users to see who else is on a page, to communicate with them, and to traverse the Web as a group. This will facilitate information searches and help foster community: people accessing the same pages are likely to be in search of the same type of information and to share similar interests. Furthermore, this work enriches the expressive capabilities of real-time electronic communication, by making it possible to insert hypertext links, sounds, and images amid normal conversational text.

- **Visual Who**: The population of a community creates many patterns. Some are patterns of activity, such as the ebb and flow of people at various times of day. Others are patterns of affiliation: the clustering of people who share similar jobs or interests. In the real world, these patterns are visible and help create a vibrant social environment. Yet in the virtual domain, they are difficult to discern. Visual Who makes the social patterns of a virtual community visible, creating a dynamic visualization of the members' affinities and animating their arrivals and departures. By allowing the user to create many different views of the community structure and to observe the temporal patterns created by the members' activity, Visual Who makes it possible for the user to see a multi-faceted overview of a complex society.

112. Hyperinstruments

(Professor Tod Machover)

The Hyperinstrument project started in 1986, with the goal of designing expanded musical instruments, using technology to give extra power and finesse to virtuosic performers. Such hyperinstruments were designed to augment guitars and keyboards, percussion and strings, and even conducting, and have been used by some of the world's foremost performers, such as Yo-Yo Ma, the Los Angeles Philharmonic, Peter Gabriel, and magicians Penn & Teller. Since 1992, the focus of the Hyperinstrument group has expanded in an attempt to build sophisticated interactive musical instruments for non-professional musicians, students, music lovers, and the general public. Systems such as Drum-Boy and Joystick Music allow non-musicians to shape and create complex and interesting musical pieces by using gestures or word descriptions (such as musical "adjectives") to influence the real-time interactive environment. Current Hyperinstrument research is attempting to push the envelope in both of these directions: by designing high-level professional systems that measure the most subtle and sophisticated human performance (such as current projects with the rock musician formerly known as Prince and classical artists like the Kronos Quartet); and by building ever-more-powerful, interactive entertainment systems for the general public (such as interactive music games, music learning systems, and Internet-oriented group performance and creation). The research focus of all this work is on designing computer systems (sensors, signal processing, and software) that measure and interpret human expression and feeling, as well as on exploring the appropriate modalities and innovative content of interactive art and entertainment environments. During 1995, we will be further expanding the Hyperinstrument system to include gestural and intuitive control of visual media.

113. Brain Opera

(Professor Tod Machover)

Today's world of networked electronic communications offers the possibility of building communities for creative expression, while new technologies allow artistic and entertainment experiences to be actively enjoyed by audiences. The Brain Opera is one of the largest scale interactive music/image projects ever undertaken, and will push forward many of the recent developments of the Hyperinstruments group. This new generation of Hyperinstruments will allow audiences to explore a web of interactive music "games," recording, shaping and producing music and images that will be incorporated into collective "performances." The music and images of the Brain Opera will always retain their essential character and impact, but will be modified and changed by each audience. The work will be premiered at New York's Lincoln Center (at their prestigious new international festival) and simultaneously on the Internet (as a main event of the Internet World's Fair) in Summer 1996. Numerous new technologies and concepts are currently being developed for this project, including new sensors for individual interaction, sensors for measuring the reaction/feeling of a large group of people (approximately 100), techniques for automatically classifying and editing audio samples and reusing them in a musical context, tools for "shaping" and "personalizing" a piece of music while it is playing, and a framework for connecting physical events in public spaces with virtual participation on the Internet.

THE VISIBLE LANGUAGE WORKSHOP

The thrust of research at the Visible Language Workshop (VLW) is toward the integration of two main directions: first, the study of the design of form and content in the new electronic information medium; second, the potential relationships among artificial intelligence, graphics, and technology. We continue to build a seamless, sophisticated graphical, dynamic, and interactive testbed to use for prototyping tools, editors, and scenarios. Simultaneously we are developing new artificial intelligence methods to help filter, prioritize, and design information on-the-fly, and to capture qualitative and intuitive design knowledge.

114. Information Landscapes

(Ronald MacNeil, David Small and Suguru Ishizaki)
In contrast to the discontinuous jumps between flat pages provided by the ubiquitous hyperlinked Web browser, an information landscape provides a continuum of paths from place to place, allowing loose and spatially meaningful relationships between information objects. Navigation of such spaces should not only reveal needed and appropriate information, but should do so in such a way that the journey itself has meaning.

- **Galaxy of News** *(with Earl Rennison):* This system embodies a scalable approach to visualizing and navigating through large quantities of independently created pieces of information, in this case news stories. It combines the effective aspects of both searching and browsing, and the ability to switch between these modes of operation seamlessly within a single interface. The system automatically organizes disconnected articles into dynamically formed groups, based on the content of the articles, and allows quick access to related information and the ability to understand the relationships among articles quickly.

- **Financial Viewpoints** *(with Lisa Strausfeld):* This project was the "virtual" manifestation of an earlier "physical" 3-D model built out of parallel planes of Plexiglas and threaded metal rods. Dynamic, transparent intersecting planes are used to represent complex information. Six mutual funds are compared on the basis of seven features such as risk and rate of return. Depending upon the way in which we approach the information—literally, the direction from which we zoom into the field of data—we can examine differently the features and the context that influence them. For example, we can retrieve straightforward numerical data from a chart to see how much a given fund rose or fell over a particular period of time. We are also able to pull up bar graphs to chart the annual rate of return against risk over the last 10 years, or look up the profiles of the men and women who direct the funds. The transparent quality of the visual structures and the variable focus of data help the viewer maintain the context. This helps prevent the feeling of being lost in a 3-D typographic space.

- **GeoSpace** *(with Ishantha Lokuge):* The visualization of complex geographic information is a challenging design problem. Our goal was to create an interactive visualization system that can determine the interest of information seekers based on their queries, and can visually guide their attention in the complex data space. We have adopted information-seeking dialogue as a fundamental model of interaction, and have implemented a prototype geographic information system that progressively provides information upon a user's queries. The user-queries trigger presentation plans composed by domain knowledge that reflect a user's interest. The presentation plans are linked by an activation spreading network which creates a highly responsive visual environment. Dynamic use of various visual design techniques, such as transparency, 3-D graphics, and typography, are integrated to enhance the clarity of the visual display.

- **Millennium Project** *(with Earl Rennison and Lisa Strausfeld):* In 1912 the S.S. Titanic sank on its maiden voyage, Woodrow Wilson won the U.S. presidential election, Sun Yat-sen founded The Kuomintang (the Chinese National Party), C.G. Jung published *The Theory of Psychoanalysis*, Edwin Bradenburger invented a process for manufacturing cellophane, and Marcel Duchamp painted "Nude Descending a Staircase." How, if at all, do these events relate to one another? Where, when, and what were the confluences of ideas and people that influenced the outcome of these events? How do we acquire the knowledge to understand the complex associations between people and ideas, across time and place, based on the artifacts and events they created?

  The Millennium Project set out to address these issues by developing a conceptual and computational approach for enabling understanding of a large, multidimensional set of information. The goal of the Millennium Project is to provide a knowledge seeker with the ability to move through virtual time and space to explore and discover the connections among artifacts of philosophy, painting, music, literature, science, and political events of a pivotal time in world history: the years from 1906 to 1918. This virtual space continually constructs and reconstructs itself based on the knowledge seeker's movements through and within it, much like the process of moving through the conceptual spaces of our minds as we construct meaning. The conceptual framework for this research is based on linguistics and cognitive science. This project shows how our concepts of embodied cognitive models and visual discourse assist us in designing and building a computational environment that enables people to understand large bodies of information.

- **Dynamic Timelines: Visualizing Historical Information in Three Dimensions** *(with Robin L. Kullberg)*: This research reinvents the traditional timeline as a dynamic, 3-D framework for the interactive presentation of historical information. An experimental visualization of the history of photography uses new visual techniques, such as infinite zoom, translucency, and animation, to present a database of over 200 annotated photographs from the collection of the George Eastman House. Dynamic, interactive design solutions address the communicative goals of allowing seamless micro- and macro-readings of information at several levels of detail and from multiple points of view.
- **Enhancing Air-Traffic Control Information** *(with Jeffrey Ventrella)*: Conventional ATC displays are dangerously confusing. This prototype creates visual analogs for such ATC concepts as "green space" (allowable distance between planes) and flight corridors in a 3-D space environment. Smooth transitions between tower, pilot, and global views maintain context. What-if scenarios dealing with weather conditions and flight paths can be replayed for training scenarios.

115. Intelligent Graphics: Meta-Design

(Ronald MacNeil and Suguru Ishizaki)
In order to design information in the age of dynamic, electronic media more effectively, we need a richer understanding of the structure of information (i.e., the architecture of information), and the tools to manipulate that structure. New tools for designers will incorporate these new structural representations to assist during construction, or later during presentation, when the designer may not have the luxury of overseeing information's design.

- **Dynamic Magical Environment** *(Ishantha Lokuge and Suguru Ishizaki)*: This project explores salient principles of close-up magic as a means of creating engaging interactive systems. Although previous research has identified the need to engage the user during the interaction process, there is a dearth of design methods for doing so. The term "engaging" is generally defined as "to embody characteristics that include responsive reaction, unobtrusive interface, guided navigation, suggestive exploration, and unexpected behavior." Principles in magic, such as framing context, focus of attention, continuity, adaptation, element of surprise, and timing and pacing, are studied and used as a guide to designing engaging interactive systems. Two experimental information systems have been implemented in the domain of geographic information: GeoSpace and MediaMagic. Both systems demonstrate the application of the principles and exemplify an approach for designing engaging interaction.

- **Temporal Typography** *(Yin Yin Wong and Suguru Ishizaki)*: Text is no longer limited to a static presentation in electronic communication. Typographic form can change in size, color, and position according to a reader's interaction in real time. This project proposes "temporal typography" as an area of study that incorporates the dynamic visual treatment as an extension of written language. Presently, graphic design lacks ways to conceptualize and describe temporal aspects of typography in a systematic and analytical fashion. We have developed a characterization scheme that provides a set of concepts and terminology that allows for the description of typographic expressions that change dynamically over time. A software tool, along with a scripting language, were implemented based on the scheme. This scheme has been examined through the creation of various design examples that explore dynamic treatment of text, including expressive e-mail, visualization of dialogue, and typographic tone of voice. Our current effort focuses on the design of visual interface that support the writing of temporally expressed text.
- **VIA** *(Dr. Louis Weitzman):* This thesis describes research to support the design and presentation of information in computer-based, multimedia documents. The "design and orderly arrangement" of information (i.e., the architecture of information), is a growing problem for future computational environments. Increases in computational power and the increased bandwidth of interconnected networks provide greater access to information. These factors, combined with the realization that not all of this information can now be pre-designed, necessitate new tools and techniques to ensure the effective presentation of computer-based information.

  This research investigates the use of relational grammars (an extension to traditional string languages) to enable the design of dynamic documents based on formal visual languages. The description of design as a formal language affords a number of different interaction styles and support tools to the process of design. This research supports design in three distinct ways. First, information given to the system is automatically presented within a predefined language and style. This dynamic design reconfigures the same information, making it sensitive to the output environment in which the documents are delivered. Second, the creation of information can be supported by improved design decisions as the design progresses. These improvements help the user explore the design space with incremental design decisions. Finally, this class of visual languages is constructed by demonstration, using separate visual languages in a meta-design authoring tool.

- **TYRO**: This is an evolving system which uses case-based reasoning for capturing and reusing media communication design knowledge. The task in each domain—such as page layout, video editing, dynamic electronic instruction manual production, or information landscapes—is to:

  • express visually in editors and browsers the object attributes and behaviors, along with their underlying representation;

  • provide a facility for demonstrating the canonical examples and capturing these in a case library; and

  • for any novel case, be able to retrieve the most applicable cases and adapt them, given an evolving set of adaptation reasoning mechanisms.

- **Typographic Performance** *(Suguru Ishizaki):* This project explores the possibility of developing visual design solutions for computational media that can continuously respond to changes in the information and the goal of its recipient. A design solution, such as a display of online news, is considered a performance composed of a number of active design agents (or performers), each of which is responsible for presenting a particular aspect of information, such as a headline and a news story. The individual design agent is sensitive to changes in its situation (the information itself, the goals of the reader, and the other design agents in the visual scene). The solution as a whole emerges from the dynamic activities of collaborating agents. This model is fundamentally different from the traditional view, which considers a design solution to be a set of affirmative statements. This proposed model suggests a new perspective on the way designers perceive a design solution: an "active entity" consisting of a collection of design agents. A multi-agent design testbed system has been implemented, along with high-quality typographic tools. We are currently developing various design examples in order to help evaluate the model as well as the computational technique.

- **Darwin Meets Disney** *(Jeffrey Ventrella):* This research shows the use of genetic algorithms in graphic design. It emphasizes expression and motion style as demonstrated by animated characters. The animated characters can evolve behaviors like walking, using automatic evolution. They can also evolve expressive style to their motions via interactive techniques from the user.

116. DataWall: Large-Scale High-Resolution Displays

(Ronald MacNeil)
A 2000 x 6000-pixel seamless hybrid made by overlapping three 200-inch 2K x 2K Sony monitors, is still the highest resolution "TV" in the world. It supports experiments in group decision-making, testing the limits of just how much information can be displayed at once. A fully dynamic, large-scale DataWall projection display is in the works.