Multimedia Research Platforms

David A. Berkley
J. Robert Ensor

Several AT&T Bell Laboratories research projects focus on multimedia information, providing new technologies for the creation, storage, transmission, retrieval, presentation, and use of such data. Other projects focus on inventing new ways to structure and use multimedia data by developing research platforms and novel applications. This paper discusses three of these platforms and associated applications: Rapport and MR, multimedia desktop conferencing and virtual meeting room service platforms; HuMaNet, a speech-controlled multitechnology conference room environment; and the Nemesis information browser, which provides desktop access to remotely stored papers, photographs, and presentations.

Introduction

Several research projects within AT&T Bell Laboratories—focusing on multimedia information—are providing new technologies for its creation, storage, transmission, retrieval, presentation, and use. Studies currently under way address such diverse topics as disk storage layout, video coding techniques for low-bit-rate cellular phone links, and signaling protocols for broadband networks. Complementing these basic technologies projects, other researchers are developing new ways to structure and use multimedia data by constructing research platforms and novel applications.

The platforms provide general environments within which researchers can create future multimedia structures and study appropriate technologies. Although many of the concepts discussed in this paper have led to development activities[1]—some using important aspects of the platforms—the focus is on research generality rather than cost or development considerations. Ideally, these platforms should allow the creation of new knowledge of multimedia technologies and architectures. This knowledge should foster the development of new paradigms for the human interaction and basic structures (software, hardware, and middleware) needed for AT&T to exploit them.

Many research platforms exist. The three discussed in this paper were chosen because collectively, they cover the major areas of importance for future AT&T multimedia technologies. Specifically, the platforms represent foundations for multimedia information exchange and services development, a collection of multimedia technologies and their evaluation, and a multimedia data-storage and retrieval system. They include:

- *Rapport and MR*, multimedia desktop conferencing and virtual meeting room service platforms;
- *HuMaNet*, a speech-controlled multitechnology conference room environment; and
- *Nemesis*, an information browser that facilitates desktop access to remotely stored papers, photographs, and presentations.

Rapport, MR, and Related Projects

The *Rapport* multimedia conferencing system was developed to learn how computers and telecommunications networks can be combined to help people communicate more effectively over long distances. Rapport permits geographically distributed users to interact with each other through real-time exchange of information in several media, including voice, video, and computer-based text, graphics, and images.[2] This system

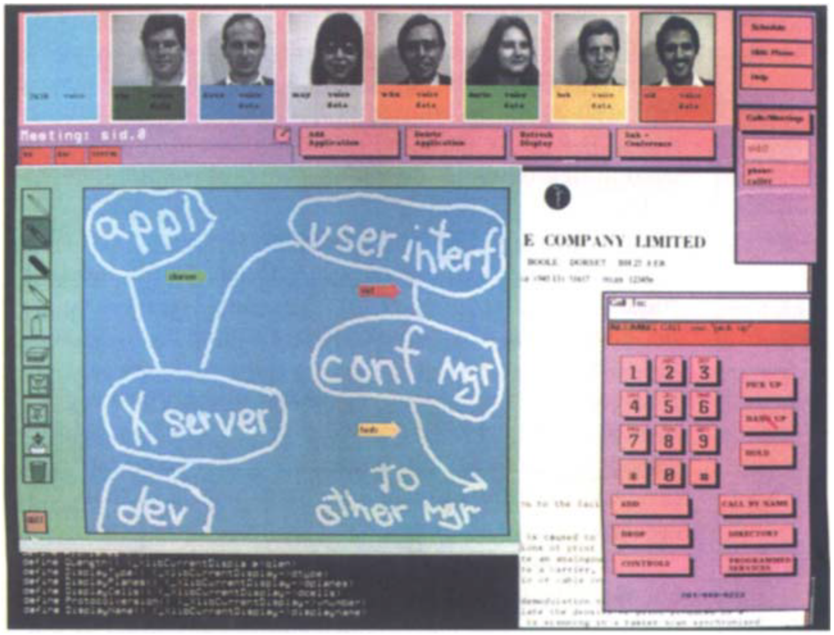


Figure 1. This composite illustration shows a computer screen during a Rapport conference. Pictures of each participant appear at the top, and all eight individuals are communicating by telephone. Seven participants share program displays and inputs through their computers. The programs *bb*, *doc*, and *xterm* are associated with the call. With a CAD tool, *bb* allows the participants to work together; *doc* helps them make simple drawings on scanned documents; and *xterm* gives them access to conventional, text-based programs. The control panel is used to place or answer calls.

coordinates computer and network resources to create an environment for communication. A goal of the Rapport project is to create environments rich enough for users to interact over long distances, just as if they were in the same room.

Rapport allows its users to participate in meetings without having to leave their offices. Users "meet" in a virtual meeting room, a computer-based electronic environment that acts as the analog of the physical space in which face-to-face meetings occur.

When "inside" a Rapport virtual meeting room, participants share information in ways similar to sharing objects within a physical room. For example, when participants enter a Rapport meeting, they can be seen and heard by other conferees just as if they had entered a common, physical room. Similarly, placing a computer program display in the virtual meeting room allows all the conferees to see it, which is analogous to using a wall monitor in a lecture hall.

By integrating a variety of computing and communication resources with consistent control and access methods, Rapport provides its users with a unified communications system. Users may access Rapport conferences by employing a variety of devices. Participants having only telephones can hear discussions, while those having interconnected computers can view and operate each other's programs.

Computer programs associated with Rapport conferences produce identical windows on the computer screens of meeting participants. Thus, computer-generated data and displays are available to the conferees, enhancing information exchange during meetings. Users having monitors are able to see video images of each other, the so-called "talking heads" representation of conference participants. Rapport also supports use of shared video within conferences. This capability supplies common video displays for participants, akin to a television broadcast for the group.

Figure 1 shows a computer screen during a Rapport conference. Along the top is a representation of each conference participant, and all eight individuals are communicating using their telephones. Seven participants share program displays and inputs through their computers. The programs *bb*, *doc*, and *xterm* have been associated with the call. With a CAD tool, *bb* allows the participants to work together; *doc* helps them make simple drawings on scanned documents; and *xterm* gives them access to conventional, text-based programs. The telephone-like control panel is used to place or answer calls.

MR. Rapport has served as the basis for other

Panel 1. Abbreviations, Acronyms, and Terms

bb—a software program that allows participants in Rapport to work together

CAD—computer-aided design

CAIP—Computer Aids to Industrial Productivity, a center at Rutgers University in New Brunswick, New Jersey

CORAL—an AT&T client/server directory service that provides indexing of multimedia information

CPU—central processing unit

doc—a software program that allows participants in Rapport to make simple drawings on scanned documents

HuMaNet—a made-up word standing for "Human/Machine/Network," which is an AT&T experimental communications system based on ISDN technology

ISDN—integrated services digital network

middleware—software that helps build and use multimedia applications

MJPEG—Motion Joint Photographic Experts Group

MPEG—Motion Picture Experts Group

MR—meeting room, an AT&T multimedia, multiparty communication system that extends the virtual meeting room concept of Rapport

NTSC—National Television Systems Committee

VMM—visual meeting minutes, an AT&T system that helps automate the recording of meeting activities

whiteboard—a computer interaction tool that is the electronic equivalent of chalk and blackboard

xterm—a software program that allows participants in Rapport to access conventional, text-based programs

research and development projects. MR (for meeting room) is a multimedia, multiparty communication system that extends the virtual meeting room concept of Rapport.[1]

In the MR system, a meeting room may exist longer than the conferences held within it. An MR meeting room provides a context in which meetings take place, and because a room may persist, offers a source of continuity for multiple communication sessions. For example, a user might set up a room before a meeting so that conferees have a specialized setting for their group activities. (More specifically, a person might associate a set of application programs with the room and arrange their displays to create a structured context for later meetings.)

Continuity of the group's activities might be enhanced by preserving the results of a meeting within a room for later use. The room can also be accessed by one person (a "meeting" with one participant) and serve as a storage and presentation environment for multimedia information. Thus, an MR system user may move seamlessly between private and group activities within this multimedia information environment. In summary, the MR server provides environments for sharing information during meetings and for persistent information storage to support long-term collaborations.

MR is a much more modular and extensible system than Rapport. As Figure 2 illustrates, the MR system is a collection of specialized servers coordinated through the MR server. The system illustrated is composed of an MR server, a name server, an event distributor, and three specialized servers, each corresponding to a medium (voice, video, and data). The figure depicts a deployment in which these servers are located within the network. It also portrays the agents (clients) who have been implemented to execute on users' workstations. In particular, a conversation manager—as well as local managers for voice, video, and data—are shown executing on two user workstations.

Middleware is software that helps people build and use multimedia applications. MR is the basis for a middleware system that supports a wide range of applications including on-line data services, home-banking, and interactive television, as well as multimedia conferencing.

Visual Meeting Minutes. An ongoing research effort, based on MR and the visual meeting room model, is known as *visual meeting minutes* (VMM).[3] The VMM system helps automate the recording of meeting activities. It records both participant interactions and their manipulation of multimedia data during conferences. VMM provides an effective way to represent the activities of collaborations in which distributed participants exchange multimedia information. It records and displays control information associated with users and meeting data in a way that lets users focus on and access the records of interest.

For example, a user who wishes to access video segments of a particular collaborator can use VMM to display any video record from the channel used by that person. The VMM approach to record keeping helps users find the historical data they want by themselves. As shown in Figure 3, VMM algorithms automatically lay out such information in a visually intuitive manner and make it available through simple icon commands.

CORAL Directory Service. Based on the MR architecture, the CORAL Multimedia Directory Service is a client/server service that provides indexing of multimedia information. Initially, CORAL was designed to serve as a directory and naming component. That goal has been extended, and now CORAL is being built to provide a framework within which users can create custom views of almost any multimedia information.

CORAL can be used as a basis for creating a variety of information services that require remote access to distributed multimedia data. Some examples include personalized address books, multimedia teleconferencing, electronic yellow pages, and home shopping catalogues.

The CORAL directory is a network-based service, thereby allowing access to common information from a variety of endpoints. This eliminates the need to maintain separate address books for separate platforms. CORAL also adds update notification events, a feature that allows clients to maintain consistent views.

Archways 3D Visualization System. Based on the MR persistent virtual meeting room, the Archways three-dimensional visualization system[4] explores new ways to represent multimedia communication environments. The system uses its knowledge bases to generate user interfaces automatically for shared communication environments.

Interfaces generated by the Archways three-dimensional visualization system are based on the following model of long-distance communication:

> When someone interacts with a remote entity, the individual remains at the same physical location but is placed in a virtual environment, which is simultaneously defined by the long-distance interaction. For example, when talking by telephone, the individuals are situated in distinct physical locations, as well as in a common virtual place that is hosting the conversation. These dual environments are not associated only with telecommunications among people; they also arise when a person or program accesses a remote service. When an Archways user is communicating with one or more remote parties, the system generates a representation of the physical location of each party and of the virtual room hosting their interaction.

Figure 4 shows an Archways visualization. Four people—Doree, John, Cati, and Sid—are perusing the Internet together by sharing a Mosaic browser. The group is looking at Cati's home page showing her stamp collection. The photograph shows a two-part view of the virtual room that has been automatically generated for Sid.

The visualization includes cues designed to provide information about each participant, their activity, and focus. For example, the red line connecting Doree's keyboard to the browser display indicates that she currently has input control of that program. Similarly, the blue arrow is a telepointer generated by John's mouse indicating the object of his attention—a stamp in the browser.

Montage Audio/Video Bridge. Montage is a hardware and software system[5] designed to provide high-quality audio and video bridging. It is used within the MR environment for sharing information in these two media.

Unlike most current systems, Montage supports multiple video streams, making possible such applications as continuous-presence video teleconferencing and digital picture-in-picture. With its multiple active video streams on display, Montage provides users with the ability to customize the placement and size of video windows, just as text and graphics windows can be manipulated.

Montage incorporates new techniques for implementing multiresolution, multistream video compression and decompression systems using generic codecs. This approach provides multiple video functionality to standard systems designed for handling

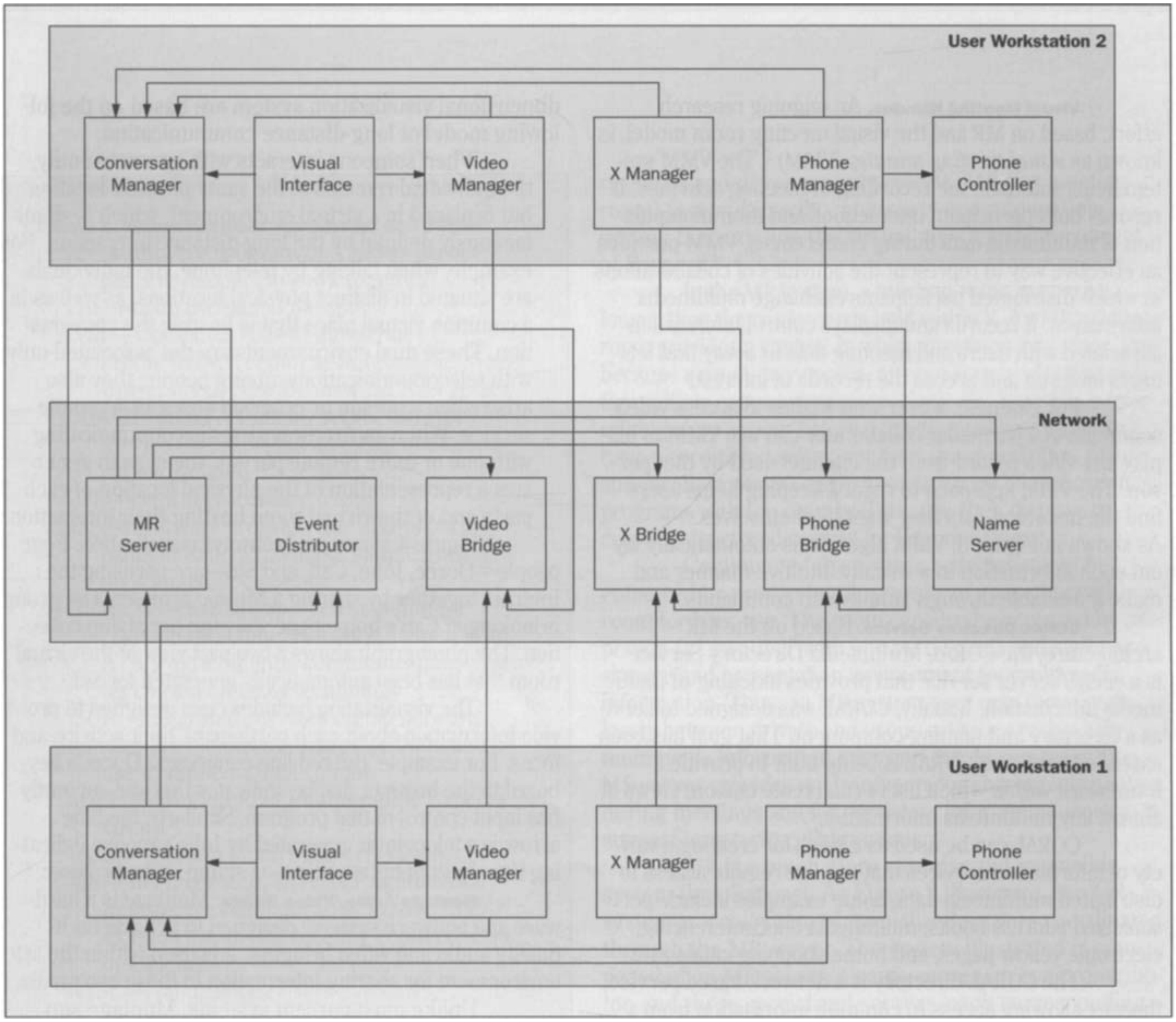


Figure 2. The MR system is a collection of specialized servers coordinated through the MR server. The system illustrated is composed of an MR server, a name server, an event distributor, and three specialized servers, each corresponding to a medium (voice, video, and data). The block diagram depicts a deployment in which these servers are located within the network. It also portrays the agents (clients) who have been implemented to execute on users' workstations. A conversation manager—as well as local managers for voice, video, and data—are shown executing on two user workstations.

single video streams, requiring almost no extra memory or processing power.

Two Montage systems have been implemented. One system, based on Motion Joint Photographic Experts Group (MJPEG) standards, uses inexpensive and readily available MJPEG chip sets to provide better compression. Another system, based on Motion Picture Experts Group 2 (MPEG-2) standards for video, uses standards-compliant video- and software-based transforms. It will, however, use MPEG-2 chip sets when they become available.

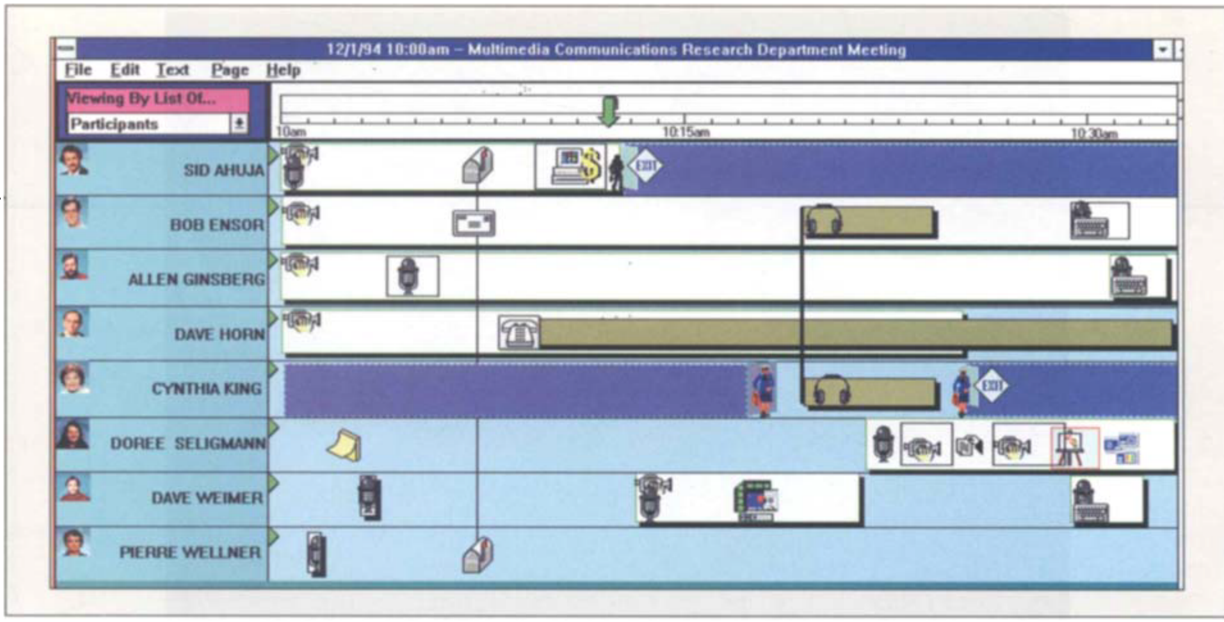


Figure 3. An ongoing research effort, known as visual meeting minutes (VMM), helps automate the recording of meeting activities. It records both participant interactions and their manipulation of multimedia data during conferences. It records and displays control information associated with users and meeting data in a way that lets users focus on and access the records of interest by themselves. As shown in the illustration, VMM algorithms automatically lay out such information in a visually intuitive manner and make it available through simple icon commands.

Business Product Protoypes. Rapport has influenced the development of several AT&T products, including Rhapsody groupware, the TeleMedia (Vistium™) multimedia platform, and WorldWorx™ Solutions. MR technology is used as a base for the AT&T Global Business Communications Services' Multimedia Call Server[1] and AT&T Business Communications Services' service prototypes—for example, the Virtual Meeting Service.

HuMaNet

The *HuMaNet* (a made-up word standing for "Human/Machine/Network") project was started to understand the interaction between human interface technologies. The technologies employed include:

- *Speech recognition*, how people can talk to machines and networks;
- *Speech synthesis*, how machines and networks can talk to people;
- *Speech and image coding*, mechanisms for efficiently representing information from various input sources, moving it through networks, storing it, and replaying it to users on demand;
- *Speaker verification*, methods that facilitate protection of valuable information stored on machines or in a network;
- *Conferencing* by means of audio and video media, allowing direct person-to-person communications in a multimedia environment; and
- *Electroacoustics*, various microphone and loudspeaker technologies that allow users to operate the complex HuMaNet system hands free.[6]

In the HuMaNet system, a computer-controlled microphone[7] picks up commands from a user. Such commands can control—through the speech recognizer[8]—local room lights; access to local databases (for example, viewgraphs) and remote databases (for example, compressed images and text read back by the text-to-speech synthesizer); and the placing of telephone or video ISDN calls. The audio pickup for telephone calls is made by means of a fixed-array microphone.

Evaluation of Multimedia Systems. One reason for constructing such a complex technical arrangement is to provide a mechanism by which researchers can gain some understanding of how the technologies might be used in the future. More importantly, this type of platform is useful for understanding areas in which complex technologies work well together and those in which they may cause difficulties or failures when used together in future multimedia systems.

Technical interactions. Some interactions of particular interest are purely technical. For example, a surprisingly good interaction was the combination of the

Figure 4. This is a photograph of an Archways visualization. The people pictured are perusing the Internet together by sharing a Mosaic browser. The group is looking at one individual's home page. A view of the virtual room that has been automatically generated is also shown. It includes cues designed to provide information about each participant, activity, and focus. The mouse-generated telepointer indicates the object of attention.

speech recognizer and computer-controlled microphone. The recognizer was speaker trained and designed to reject commands not in its 80-word vocabulary and not used with proper grammar. (Millions of allowed commands for the HuMaNet system currently exist, not counting valid telephone numbers, all of which are also included in the grammar.)

The microphone array forms an acoustic beam pointed at the current talker. The acoustic beam isolates the talker from other potential sources of noise in the conference room (see Figure 5)—including other people—while allowing the talker to move freely around the room.

This microphone system provides interrupt protection against all but a determined interferer. It was originally designed as part of a family of conferencing microphones. Combined with the speech recognizer, it was one

Figure 5. In the HuMaNet system, a microphone array, shown above the video screens at the front of the table, forms an acoustic beam pointed at the current talker. The acoustic beam isolates the talker from other potential sources of noise in the conference room while allowing the talker to move around the room freely.

of the first applications in which speech recognition could be carried out across a room with reasonable accuracy.

Since then, array microphones have been studied for applications in a variety of distant talker, speech recognition systems, such as drive-up kiosks. Additionally, they have been used in Disney Enterprises' *Alice in Wonderland* system, an interactive video environment allowing a user to give voice commands that control the video display.[9]

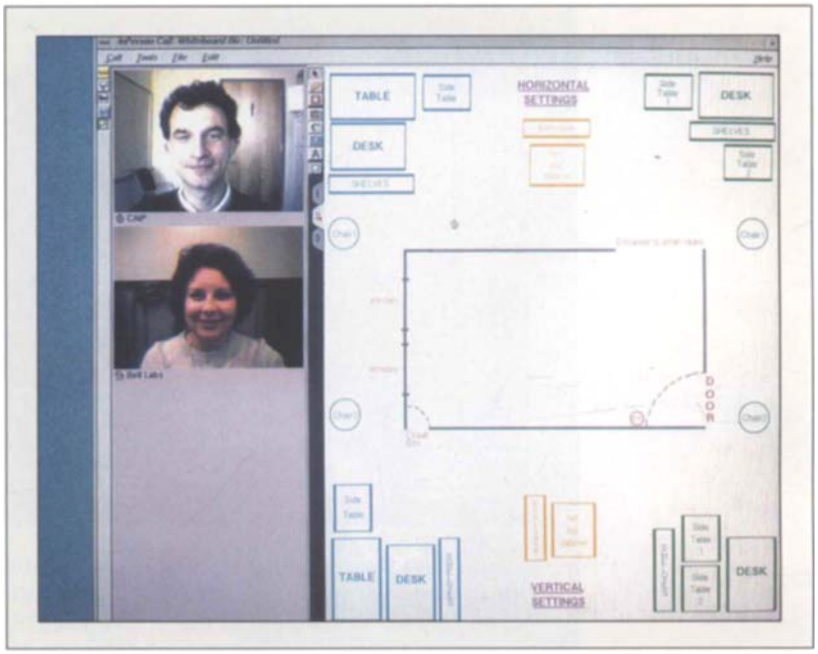
Sometimes, however, these complex technical systems have unpleasant surprises in store for researchers. Examples of this are the apparently simple use of voice response (text-to-speech synthesis in HuMaNet) for confirming commands and for a variety of other purposes, such as reading back textual information stored in hyper-media databases.

This latter application emphasized the need to enable a user to "talk over" or "barge in" during incoming speech to invoke additional voice commands. This situation was experienced in some early network applications of speech recognition. However, barge-in is far more difficult to achieve when combined with hands-free operation.

The problem here is similar to that encountered in providing full-duplex audio conferencing using acoustic echo cancellation. An acoustic echo canceler creates an electrical analog model of the room in which it is situated and subtracts a calculated echo from the signal, plus real echo, returning from the room.[10] Using this technology would allow barging in when the feature is combined with the computer controlled microphone used for speech recognition. Such a combination is still the subject of ongoing research. (Acoustic echo cancellation is used in HuMaNet for video telephony. In this instance, a fixed microphone array is used and acoustic echo cancellation is relatively straightforward.)

Human factors. The interaction of human users with such complex technological systems is an equally important and even less understood aspect of multimedia systems. Speech recognition and speaker verification were incorporated into HuMaNet because it was believed this would make the system easier to use. Anecdotal reports make this believable, but a formal study has not been carried out. In fact, a fundamental question still remains: *How important is any given aspect of a multimedia system for the user?*

Figure 6. As part of the HuMaNet project, a series of experiments was conducted at AT&T Bell Laboratories in Murray Hill, New Jersey, to study a complete multimedia communications system. Participating in the experiment were 64 teams of two people. Each team experienced one of the experimental conditions and participated in both scenarios of system operation, which included a free conversational task and a task aimed at arranging furniture in a joint hypothetical future office. A simple head-and-shoulders video conference was held, and the whiteboard was essential to the furniture arranging task.



For example, consider video telephony—especially recent systems that combine a video connection with some form of computer interaction, such as a *whiteboard.* Historically, if a formal evaluation of the video system was done, results would indicate that people are very critical of video quality when the video is viewed in a passive mode—that is, not part of a communications system.

Modern, low-bandwidth video telephony systems—when tested in this one-way passive mode—tend to receive scores of three or lower on a scale of one to five, with one denoting unsatisfactory performance and five excellent performance. (Such results have been obtained in a variety of experiments conducted at AT&T Bell Laboratories and in collaboration with the AT&T Consumer Laboratory in Murray Hill, New Jersey.)

Recently, a study of a complete multimedia communications system was undertaken. The study included audio/video and whiteboard in a realistic communications application. The system used a commercial video communications system in which the quality was controllable by changing both the bit rate and frame rate. (The frame rates used were 5 and 20 frames/s.) A wideband audio path was available (uncompressed 16-bit pulse-code modulated, sampled at 16 kHz, resulting in a data rate of 512 kb/s, matched to the delay incurred by the video path (about 0.3 s). However, the audio quality was regulated explicitly by controlling the audio bandwidth.

The bandwidths used in the experiment were 5 kHz and 2 kHz. The audio/video system also incorporated a whiteboard that could be accessed simultaneously from both ends of a two-point connection. The maximum total bandwidth used by the system was 2,155 kb/s.

The connection was made using XUNET-II,[11] an experimental nationwide AT&T network linking various AT&T and university sites throughout the United States with 45 Mb/s or greater asynchronous transfer mode digital service. The network connected one of the HuMaNet laboratories at AT&T Bell Laboratories in Murray Hill, New Jersey, and the Computer Aids to Industrial Productivity (CAIP) Center at Rutgers University in New Brunswick, New Jersey, about 15 miles away. Thus, realistic tasks could be done by users located some distance from each other.

The system was configured for completely hands-free operation. Because of the video-path delay mentioned earlier, acoustic echo cancellation had to be provided to avoid echoes or unacceptable clipping of speech. Development models of the AT&T Custom Electronic Systems' QuiteQuiet™ acoustic echo canceler were used in conjunction with good quality microphone and loudspeaker arrangements.[12]

Participating in the experiment were 128 people (64 teams of two, with one individual at AT&T Bell Laboratories and the other at The CAIP Center) who generally were not experienced computer or video conferencing users. Each team experienced one of the experimental conditions (that is, a combination of video frame rate and audio bandwidth) and participated in both scenarios of system operation, which included a free conversational task and a task aimed at arranging furniture in a joint hypothetical future office. Conversation took place as a simple head-and-shoulders video conference. The whiteboard was essential to the furniture arranging task (see Figure 6).

The experimental design was balanced across many variables. In addition to observation, extensive questionnaires were completed after each phase of the experiment. Many interactions—some quite complex—are observed in the final data analysis. However, there is one very important result: the overall rating of the communications system for almost all conditions was greater than 4.0 (good) on the five-point scale mentioned previously.

This result underscores the point made earlier—that is, the study of individual multimedia system components can be very misleading. In a well-designed system, the combination of parts can lead to a satisfying user experience even when it is known that the individual parts might not merit a satisfactory rating. It is certainly true that the obverse can also easily occur. For example, a system in which each individual part merits an excellent rating may result in unworkable or unsatisfactory systems when combined for a specific customer purpose.

Understanding how people interact with such complex multimedia systems is vital for the future. Studies of specific examples of such systems will continue. In addition, the tools used in conducting such studies will be improved.

Nemesis

To study the ways in which to store and retrieve multimedia data over networks, the Nemesis network multimedia information system was built.[13] Nemesis manages a database containing talks given at AT&T and includes accompanying documents and viewgraphs.

Such World Wide Web browsers as the National Center for Supercomputing Activities' Mosaic and the Netscape Communications Corporation's Netscape Navigator* can be used to select items of interest from a Nemesis server. When a talk is selected, the compressed video, audio, and viewgraphs are sent to the requesting computer for decompression. A major difficulty with viewing such talks on National Television Systems Committee (NTSC) video is that the viewgraphs are poorly reproduced.

To solve this problem, Nemesis processes viewgraphs, handouts, and other printed material separately and displays them at high resolution by using the Ferret document system.[14] Ferret provides fast access to material compressed by the G4 facsimile algorithm. The video and viewgraphs are seen concurrently, and hypermedia links synchronize presentations.

Nemesis addresses several problems that are associated with delivering continuous media over variable-bandwidth networks and implements effective protocols for degrading service in times of network congestion. The Nemesis file-access protocols are receiver controlled. They make use of scaling and dilation to cope with network congestion and contention for CPU cycles in the receiving workstation.

Scaling is performed by dropping video frames, and a predictive prefetch algorithm is used to fetch in advance only those video frames likely to be transferred and decompressed in a timely fashion. When congestion becomes so great that even no video transfer causes overload, Nemesis dilates audio playback—without changing the sound pitch—to reduce the bandwidth requirements further.

AT&T has made Nemesis available on the World Wide Web since June 1993. Many hours of material are available on line, and these electronic archives have been viewed by more than 1,500 people worldwide.

User feedback has been very favorable. Individuals especially appreciate the ability to scan rapidly through presentations, using viewgraphs as indices. Users have found that most information in a talk is contained in the speech and on the viewgraphs. Thus, video-based information in these archives is less important, but it does help to maintain user interest and attention.

Conclusion

The ability to manage multimedia information in huge quantities is the key to the information age. The combination of a wide variety of technologies is the critical infrastructure. However, the integration of these technologies into applications and platforms (both for research and for building multimedia products) is the critical step in providing users with habitable mechanisms for using the manifold technologies.

The research platforms discussed in this paper provide a glimpse of the ways in which to use multimedia information and communications in the future. These platforms integrate advanced technologies into applications and pseudoapplications that allow studying the interaction of the technologies involved. Such interaction features close to real-world but controllable conditions and real users.

Using the information gathered in the studies will drive both future technologies and their applications toward the goal of an information-rich world in which users can access the information naturally (for example, by means of voice interfaces) and transparently. Ease of use and human-centered design—driven by research on realistic multimedia platforms—will make this scenario possible.

Acknowledgment

The authors gratefully acknowledge the contributions of AT&T Bell Laboratories associates Elisabeth de Carvalho, Darren A. Kall, and Kathleen L. Shipley in allowing the use of the unpublished results of their experiments—part of the HuMaNet project—to describe evaluation of a complete multimedia communications system as detailed earlier in this paper.

*Trademark

Netscape Navigator is a trademark of Netscape Communications Corporation.

References

1. See, for example, S. R. Ahuja, H. F Haisch, R. S Ramamurthy, and L. M. Sanders, "Multimedia Collaboration," *AT&T Technical Journal*, Vol. 74, No. 5, September/October 1995, pp. 46-53.
2. J. R. Ensor, S. R. Ahuja, and D. Horn, "The Rapport Multimedia Conferencing System," *Proceedings of the Conference on Office Information Systems*, Palo Alto, California, March 1988, pp. 1-8.
3. A. B. Ginsberg and S. R. Ahuja, "Automating Envisionment of Virtual Meeting Room Histories," *Proceedings of ACM Multimedia*, San Francisco, California, November 1995.
4. D. D. Seligmann, R. T. Mercuri, and J. T. Edmark, "Providing Assurances in a Multimedia Environment," *Proceedings of ACM SIGCHI*, Denver, Colorado, May 1995, pp. 250-256.
5. R. D. Gaglianello and G. L. Cash, "Montage: Continuous Presence Teleconferencing Utilizing Compressed Domain Video Bridging," *Proceedings of ICC*, June 1995, Seattle, Washington, pp. 573-581.
6. D. A. Berkley and J. L. Flanagan, "HuMaNet: An Experimental Human-Machine Communications Network Based on ISDN Wideband Audio," *AT&T Technical Journal*, Vol. 69, No. 5, September/October 1990, pp. 87-99.
7. J. L. Flanagan, D. A. Berkley, G. W. Elko, J. E. West, and M. M. Sondhi, "Autodirective Microphone Systems," *Acustica*, Vol. 73, No. 2, February 1991, pp. 58-71.
8. L. R. Rabiner, "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition," *Proceedings of the IEEE*, Vol. 77, No. 2, February 1989, pp. 257-286.
9. W. R. Porter, "Acoustic Measurements to Enable Speech Recognition Applications," *Proceedings of the Conference on Advanced Speech and Application Technologies*, San Francisco, California, April 1995.
10. J. C. Baumhauer, Jr., S. H. Early, J. H. Fikus, S. L. Gay, and M. A. Zuniga, "Audio Technology Used in AT&T's Terminal Equipment," *AT&T Technical Journal*, Vol. 74, No. 2, March/April 1995, pp. 57-70.
11. A. G. Fraser, C. R. Kalmanek, A. E. Kaplan, W. T. Marshall, and R. C. Restrick, "XUNET II: A Nationwide Test Bed in High-Speed Networking," *Proceedings of the 11th Annual Conference of the IEEE Computer and Communications Societies (INFOCOM)*, Florence, Italy, May 1992, pp. 582-589.
12. QuiteQuiet™ Acoustic Echo Cancellation Software Module For the AT&T DSP 1604, *AT&T Product Brief*, Vol. 10, No. 343, AT&T Custom Electronic Systems, Arlington, Virginia, February 1995, p. 21(1).
13. H. P. Katseff and B. S. Robinson, "Predictive Prefetch in the Nemesis Multimedia Information Service," *Proceedings of ACM Multimedia*, San Francisco, California, October 1994, pp. 201-206.
14. H. P. Katseff and T. B. London, "The Ferret Document Browser," *Proceedings of USENIX Summer Technical Conference*, Cincinnati, Ohio, June 1993, pp. 101-110.

(Manuscript approved August 1995)

David A. Berkley *is head of the Acoustics Research Department at AT&T Bell Laboratories in Murray Hill, New Jersey. He conducts research on electroacoustics, acoustics signal processing, and technology integration. Mr. Berkley holds B.E.E. and Ph.D. degrees in applied physics from Cornell University in Ithaca, New York. He joined AT&T Bell Laboratories in 1968.*

J. Robert Ensor *is a distinguished member of technical staff in the Multimedia Communication Research Department at AT&T Bell Laboratories, Holmdel, New Jersey. He conducts research in multimedia communication services. Mr. Ensor holds a B.S. degree in biology from Furman University in Greenville, South Carolina, as well as M.S. and Ph.D. degrees in computer science from the State University of New York in Stony Brook. He joined AT&T Bell Laboratories in 1981.*