

Edward A. Fox
Guest Editor

Virtual Video Editing in Interactive Multimedia Applications

Drawing examples from four interrelated sets of multimedia tools and applications under development at MIT, the authors examine the role of digitized video in the areas of entertainment, learning, research, and communication.

Wendy E. Mackay and Glorianna Davenport

Early experiments in interactive video included surrogate travel, training, electronic books, point-of-purchase sales, and arcade game scenarios. Granularity, interruptability, and limited look ahead were quickly identified as generic attributes of the medium [1]. Most early applications restricted the user's interaction with the video to traveling along paths predetermined by the author of the program. Recent work has favored a more constructivist approach, increasing the level of interactivity by allowing users to build, annotate, and modify their own environments.

Today's multitasking workstations can digitize and display video in real-time in one or more windows on the screen. Users can quickly change their level of interaction from passively watching a movie or the network news to actively controlling a remote camera and sending the output to colleagues at another location [5]. In this environment, video becomes an information stream, a data type that can be tagged and edited, analyzed and annotated.

This article explores how principles and techniques of user-controlled video editing have been integrated into four multimedia environments. The goal of the authors is to explain in each case how the assumptions embedded in particular applications have shaped a set of tools for building constructivist environments, and to comment on how the evolution of a compressed digital

video data format might affect these kinds of information environments in the future.

ANALOG VIDEO EDITING

One of the most salient aspects of interactive video applications is the ability of the programmer or the viewer to reconfigure [10] video playback, preferably in real time. The user must be able to order video sequences and the system must be able to remember and display them, even if they are not physically adjacent to each other. It is useful to briefly review the process of traditional analog video editing in order to understand both its influence on computer-based video editing tools and why it is so important to provide virtual editing capabilities for interactive multimedia applications.

A video professional uses one or more *source* videotape decks to select a desired video shot or segment, which is then recorded onto a *destination* deck. The defined *in* and *out* points of this segment represent the granularity at which a movie or television program is assembled. At any given point in the process, the editor may work at the shot, sequence, or scene level. Edit controllers provide a variable speed shuttle knob that allows the editor to easily position the videotape at the right frame while concentrating on the picture or sound. Desired edits are placed into a list and referenced by SMPTE time code numbers, which specify a location on a videotape. A few advanced systems also offer special features such as iconic representation of shots, transcript follow, and digital sound stores.

The combined technology of an analog video signal and magnetic tape presents limitations that plague editors. Video editing is a very slow process, far slower

UNIX is a trademark of AT&T Bell Laboratories.
MicroVAX is a trademark of Digital Equipment Corporation.
PC/RT is a trademark of IEM.
Parallax is a trademark of Parallax, Inc.

than the creative decision-making process of selecting the next shot. The process is also totally linear. Some more advanced (and expensive) video editing systems such as Editflex, Montage, and Editdroid allow editors to preview multiple edits from different videotapes or videodiscs before actually recording them. This concept of virtual viewing was first implemented for an entirely different purpose in the Aspen Interactive Video project of 1980 [4], almost four years before Montage and Editdroid were born.

Some features of analog video editing tools have proven useful for computer-based video editing applications. However, the creation of applications that provide end users with editing control has introduced a number of new requirements. Tool designers must decide who will have editing control, when it will be available, and how quickly it must be adjusted to the user's requirements. Some applications give an author or programmer sole editing control; the end user simply views and interacts with the results. In others, the author may make the first cut but not create any particular interactive scenario; the user can explore and annotate within the database as desired. Groups of users may work collaboratively, sharing the task of creation and editing equally. The computer may also become a participant and may modify information presentation based on the content of the video as well as on the user's behavior.

Different perspectives in visual design also affect editing decisions. For example, an English-speaking graphic designer who lays out print in books knows that the eye starts in the upper left-hand corner and travels from left to right and downward in columns. A movie director organizes scenes according to rules designed for a large, wide image projected on a distant movie screen. Subjects are rarely centered in the middle, but often move diagonally across the screen. Video producers learn that the center of the television screen is the "hot spot" and that the eye will be drawn there.

In each case, the information designer is visually literate, but takes for granted conventions that are not recognized by the others. Conflicts arise from the clash in assumptions. Although new conventions will probably become established over time, visual designers will for now need tools to suggest reasonable layouts and users need to be able to override those suggestions based on current needs.

MULTIMEDIA TOOLS AND APPLICATIONS

Assumptions about the nature of video and of interactivity deeply affect the ways in which people think about and use interactive video technology. We are all greatly influenced by television, and may be tempted to place video into that somewhat limited broadcast-based view. Notions of interactivity tend to extend about as far as the buttons on a VCR: stop, start, slow motion, etc. However, different applications make different demands on the kinds and level of interactivity:

- Documentary film makers want to choose an optimal ordering of video segments.
- Educational software designers want to create interactive learning environments for students to explore.
- Human factors researchers want to isolate events, identify patterns, and summarize their data to illustrate theories and phenomena.
- Users of on-line communication systems want to share and modify messages, enabling the exchange of research data, educational software, and other kinds of information.

The authors have been involved in the development of tools and applications in each of these areas, and have struggled with building an integrated environment that includes digitized video in each. This work has been influenced by other work, both at MIT and at other institutions, and the specific tools and applications have influenced each other.

Table I summarizes four sets of multimedia tools and

TABLE I. Multimedia Tools and Applications under Development at MIT

Research Area	Tool	Application	Video Editing Issues
Interactive Documentaries	Film/Video Interactive Viewing & Editing Tool	A City in Transition: New Orleans 1983-86	Recombination of video segments Seamless editing Database of shots and edit-lists Iconic representation of shots Display of associated annotation
Learning Environments	Athena Muse: A Multimedia Construction Set	Navigation Project French Language	Synchronization of different media User control of dimensions
User Interface Research	EVA: Experimental Video Annotator	Experiment in Intelligent Tutoring	Live annotation of video Multimedia annotation Rule-based analysis of video
Multimedia Communication	Pygmalion: Multimedia Message System	Neuroanatomy Research Database	Networked video editing Sharing interactive video data

applications under development at MIT's Media Laboratory and Project Athena¹ and successively identifies video editing issues raised by each. Most of this work has been developed on a distributed network of visual workstations at MIT.² This table is not intended to be exhaustive; rather, it is intended to illustrate the different roles that digitized video can play and trace how these applications and tools have positively influenced each other.

INTERACTIVE DOCUMENTARIES

Movie-making is highly interactive during the making, but quite intentionally minimizes interaction on the part of the audience in order to allow viewers to enter a state of reverie. To a filmmaker, editing means shaping a cinematic narrative. In the case of narrative films, predefined shooting constraints, usually detailed in a storyboard, result in multiple takes of the same action; the editor selects the best takes and adjusts the length of shots and scenes to enhance dramatic pacing. Editor's logs are an integral aid in the process. Music and special sound effects are added to enhance the cinematic experience.

Good documentary on the other hand tries to engage the viewer in an exploration. The story is sculpted first by the filmmaker on the fly during shooting. As all shots are unique, editing involves further sculpting by selecting which shots reflect the most interesting and coherent aspects of the story. Editors must balance what is available against exactly what has already been included and selecting those shots that will make the sequence at hand as powerful as possible. A major adjustment in one sequence may require the filmmaker to make dozens of other adjustments.

A City in Transition: New Orleans, 1983-86

The premise of the case study, "A City in Transition: New Orleans, 1983-86," was that cinema in combination with text could provide insights into the people, their power and the process through which they affect urban change. Produced by Glorianna Davenport with cinematography by Richard Leacock, the 2-hour 40-minute film was edited for both linear and interactive viewing [2]. The interactive version of this project was developed as a curriculum resource for students in urban planning and political science. Rather than creating a thoroughly scripted interactive experience, faculty who use this videodisc set design a problem that motivates the student as explorer. The editor's log is replaced with a database that contains text annotation

and ancillary documentation which are critical for in-depth analysis and interpretation.

An Interactive Video Viewing and Editing Tool

A generic interactive viewing and editing tool was developed to let viewers browse, search, select, and annotate specific portions of this or other videodisc movies. Each movie has a master database of shots, sequences, and scenes; a master "postage stamp" icon is associated with each entry. Authors or users can query existing databases for related scenes or ancillary documentation at any point during the session or they can build new databases and cross-reference them to the main database. The Galatea videodisc server, designed by Daniel Applebaum at MIT, permits remote access to the video. The current prototypes include four to six videodisc players and a full audio break-away crosspoint switcher; for some applications this is sufficient to achieve seamless playback of the video.

A viewer can watch the movie at his or her own pace by controlling a spring-loaded mouse on a shuttle bar or pressing a play button. The user can pause at any point to select and mark a segment. Segments can be saved by storing a representative postage stamp icon on a palette; the icon can later be moved to an active edit strip where it is sequenced for playback. Once a shot has been defined, the user can annotate it in a number of associated databases. A graphical representation of the shot or edit strip allows advanced users to mark audio and video separately (Figure 1).

The traditional concept of edit-list management is used to save and update lists made by users as they fill a palette or make an edit strip. Users can give each new list a name and a master icon—a verbal and visual memory aid; these are then used to place sequenced lists into longer edits or into multimedia documents. The database design significantly expands the information available to both the computer and the editor about a particular shot.

Future Directions

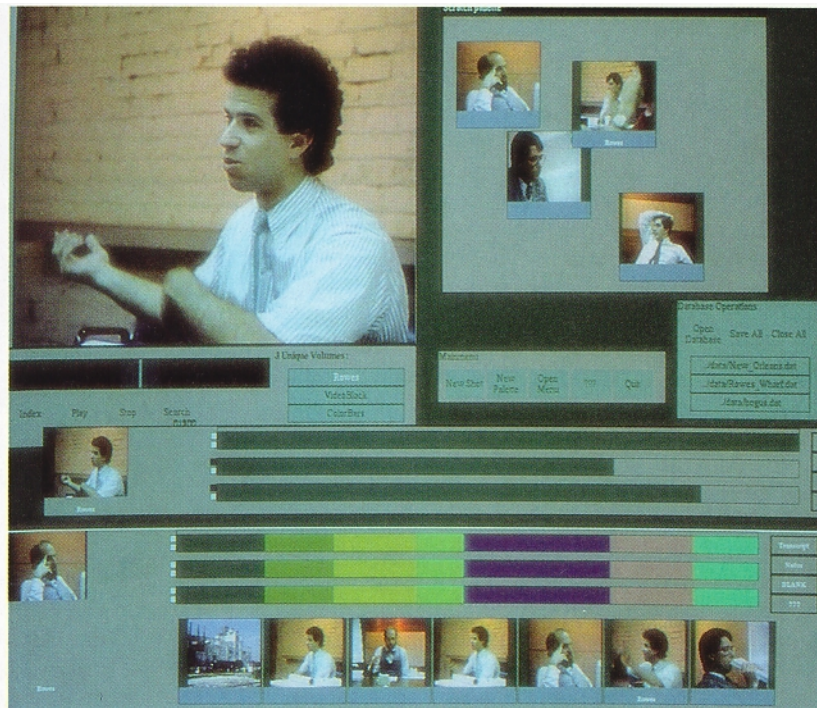
The goal of this experiment is to represent shots and abstract story models in a way that allows the computer to make choices about what the viewer would like to see next and create a cohesive story. The kind of information needed to make complex editing decisions may be roughly divided into two categories: (1) content (who, what, when, where, why); and (2) aesthetics (camera position relative to object position in a time continuum) [9].

Much of the information that is now entered into a database manually could be encoded during shooting or extracted from the pictures using advanced signal processing techniques. Digital video will also allow the computer to generate new views of a scene from spatially encoded video data. Finally, it will become easier to mix computer graphics with real images, which will both encourage the creation of new constructivist environments and make all video suspect.

¹ The Media Lab at MIT explores new information technologies, connecting advanced scientific research with innovative applications for communications media. Project Athena is an eight-year, \$100-million experiment in education at MIT, co-sponsored by Digital Equipment Corporation and IBM.

² The hardware consists of DEC MicroVAX or IBM PC/RT workstations running UNIX and the X Window System. Each workstation has a Parallax board which digitizes video in real-time from any NTSC video source, such as a video camera or a videodisc. Motion or still video can be mapped into windows on a high-resolution color graphics monitor in conjunction with text and graphics.

FIGURE 1. Film/Video Tool for Editing Video with Two Sound Tracks (Graphical Interface Designed by Hal Birkeland, MIT)



INTERACTIVE LEARNING ENVIRONMENTS

The use of digitized video in education spans a range of educational philosophies from goal-oriented tutorials to open-ended explorations. The underlying philosophy tends to dictate the level of interactivity and video editing control given to authors and users of any program. Programmed instruction and its successors are interactive in the sense that a student must respond to the information presented; it is not possible to be a passive observer. However, only the author has flexible video editing capabilities; the student is expected to work within the structures provided, following the designated paths to reach criterion levels of mastery of the information. Hypermedia provides the user with a wider range of opportunities to explore, by following links within a network of information nodes. An even richer form of interactivity allows users to actively construct their environments, not just follow or even explore. The constructivist approach provides students with some of the same kinds of editing and annotation tools as authors of applications.

The Navigation Learning Environment

In 1983, Wendy Mackay, working with members of the Educational Services Research and Development Group at Digital Equipment Corporation, began a set of research projects in multimedia educational software. The goals were to compare different instructional strategies, improve the software development process and address the technical problems of creating multimedia

object-oriented databases and learning environments [6]. Coastal navigation was chosen as a test application to push the limits of the technology. Not only does it require real-time handling of a complex set of real images, symbols, text, and graphics, but it has also been presented to students using a wide range of educational philosophies, ranging from structured military training to open-ended experiential learning at Outward Bound.

The heart of the navigation information database is a videodisc containing over 20 discrete types of information, including nautical charts, aerial views, tide tables, navigation instruments, graphs and other reference materials. The videodisc also contains over 10,000 still photographs taken systematically from a boat in Penobscot Bay, Maine, to enable a form of surrogate travel similar to the MIT Aspen project mentioned earlier.

The synchronization of images in three-dimensional space was essential to the visualization of this application. The project leader, Matt Hodges, brought the ideas and videodiscs to the Visual Computing Group at Project Athena. This project became one of the inspirations for the development of Athena Muse, a collaborative effort with Russ Sasnett and Mark Ackerman [3].

Athena Muse

At Project Athena, the required spatial dimensions for the Navigation disk were generalized to include temporal and other dimensions. In particular, several foreign language videodiscs funded by the Annenberg

Foundation were being produced under the direction of Janet Murray. An important goal was to provide cultural context in addition to practice with grammar and vocabulary. Students were presented with interactive scenarios featuring native speakers. In order to respond correctly, students needed to understand the speakers. Thus, they needed subtitles synchronized to the video and the ability to control them together.

The concept of user-controllable dimensions was created as a general solution to the control of spatially organized material (as in the Navigation project) and temporally organized material (as in the foreign language projects). Athena Muse packages text, graphics, and video information together and allows them to be linked in a directed graph format or operated independently (Figure 2). Different media can be linked to any number of dimensions which can then be controlled by the student or end user.

When reimplemented in Athena Muse, the Navigation Learning Environment used seven dimensions to simulate the movement of a boat. Two dimensions represent the boat's position on the water and two more represent the boat's heading and speed. A fifth tracks the user's viewing angle and the sixth and seventh manage a simulated compass that can be positioned anywhere on the screen. The user can move freely within the environment and use the tools available to a sailor to check location (charts, compass, looking in all directions around the boat) to set a course. Other aspects of a simulation can be added, such as other boats, weather conditions, uncharted rocks, etc. Here, the user does not change the underlying structure of the information, since it is based on constraints in the real world, but he or she can move freely, ask questions, and save information in the form of notes and annotations for future use.

Future Directions

Direct recording of compressed digital video data may provide more sophisticated ways to embed visual data into simulations of real and artificial environments. Issues of synchronization will change, as some kinds of data are encoded as part of the signal. For example, subtitles or sign language for the hearing impaired will be an integral part of the original video, obviating the need for external synchronization. Other kinds of data will continue to require externally-controlled synchronization, to handle the changing relationships among data in different applications. Users should have more sophisticated methods of controlling these associated data types, either in real time or under program control.

Cross-referencing of visual information across databases will also be easier. For example, a photograph of an island, taken from a known location in the bay, must currently be linked by hand to the other forms of information (charts, aerial views, software simulations) from which it might also be viewed. Digitally encoded representations of these images will make it easier to

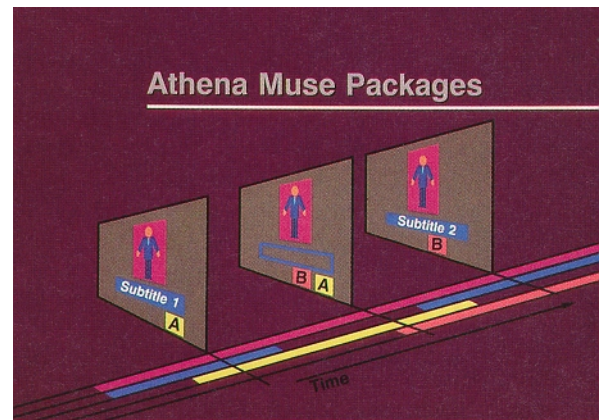


FIGURE 2. An Example from Athena Muse in which Video Segments and Subtitles are Synchronized along a Temporal Dimension

calculate these relationships.

Digital encoding would also allow views to be created that were not actually photographed. For example, the eight still images representing a 360-degree view from the water could be converted into a moving video sequence in which the user appears to be slowly turning in a circle. If enough visual information has been encoded, it may also be possible to create the illusion of moving in any direction across the water. Digital representation will also make it easier to provide smooth zooming in and out to view closeups of particular objects.

Given advances in limited domain natural language recognition and natural language parsing, it may be possible to automatically encode the audio portion of a video sequence. Future versions of the foreign language projects would then allow students to engage in more sophisticated dialogs.

VIDEO DATA ANALYSIS

Video has become an increasingly prevalent form of data for social scientists and other researchers, requiring both quantitative and qualitative analysis. The term *video data* can have several meanings. To a programmer, video data is an information coding scheme, like ASCII text or bitmapped graphics. To a researcher, who uses video in field studies or to record experiments, video data is the content, rather than the format, of the video.

The requirements researchers place upon video editing go far beyond the capabilities of traditional analog editing equipment. They want to flexibly annotate video and redisplay video segments on the fly. They often need to synchronize video with other kinds of data, such as tracks of eye movements or keystroke logs. They want methods for exploring their data at different levels of granularity, identifying patterns through recombination or compression and summarizing it for other researchers.

A clear priority for many researchers is to reduce the total amount of viewing time required for a particular type of analysis. Researchers who use multiple cameras have an even more difficult problem; they must either synchronize multiple video streams and view them together or significantly increase viewing times. Some researchers must also share control of their data, maintaining the ability to create and modify individual annotations without affecting the source data. Just as with numerical data, different researchers should be able to perform different kinds of analyses on the same data and produce different kinds of results.

The ability to integrate video and computers has spurred the development of new tools to help researchers record, analyze and present the latter form of video data. (A number of these tools are described in [11].) Wendy Mackay created a tool, EVA [8], written in Athena Muse to help analyze video data from an experiment on intelligent tutoring. The goal was to allow the researcher to annotate video data in meaningful ways and use existing computer-based tools to help analyze the data.

EVA: An Experimental Video Annotator

EVA, or Experimental Video Annotator, allows researchers to create their own labels and annotation symbols prior to a session and permits live annotation of video during an experiment. In a typical session, the researcher begins by creating software buttons to tag particular events, such as successful interactions or serious misunderstandings (Figure 3). During a session, a subject sits in front of a workstation and begins to use a new software package. A video camera is directed at the subject's face and a log of the keystrokes may be saved. The researcher sits at the visual workstation and live video from the camera appears in a window on the screen. Another window displays the subject's screen and an additional window is available for taking notes. The researcher has several controls available throughout the session. One is a general time stamp button, which the researcher presses whenever an interesting but unanticipated event occurs. The rest are buttons that the researcher created prior to the session.

Annotation during the session saves time later when the researcher is ready to review and analyze the data. The researcher can quickly review the highlights of the previous session and mark them for more detailed analysis. The original tags may be modified and new ones created as needed. Tags can be symbolic descriptions of events, recorded patterns of keystrokes, visual images (single-frame snapshots from the video), patterns of text (from a transcription of the audio track), or clock times or frame numbers. Tags that refer to the events in different processes, such as the text of the audio and the corresponding video images, can be synchronized and addressed together.

Note that annotation of live events, while useful, requires intense concentration. The mechanics of taking notes may cause the researcher to miss important

events, and events will often be tagged several seconds after they occur. Subsequent passes are almost always necessary to create precise annotations of events. While EVA does not address all of the general problems of protocol analysis, it does provide the researcher with more meaningful methods for analyzing video data.

Future Directions

Digital video offers possibilities for new kinds of data analysis, both in the discovery of patterns across video segments and in the understanding of events within a segment. Video data can be compressed and summarized in order to provide a shortened version of what occurred, to tell a story about typical events, to highlight unusual events or to present collections of interesting observations. The use of highlights can either concisely summarize a session or completely misrepresent it. Just as with quantitative data, it is important to balance presentation of the unusual data points (outliers) with typical data points. In statistics, the field of *exploratory data analysis* provides rigorous methods for exploring and seeking out patterns in quantitative data. These techniques may be applied profitably here.

An application that requires graphic overlays over sections of video needs a method of identifying the video frame or frames, a method for storing the overlay and method for storing the combination of the two. Other kinds of annotation could permit later redisplay of the video segments under program control. For example, linkages with text scripts would enable a program to present all instances in which the subject said "aha!" Researchers could use rules for accessing video segments with particular fields; for example, all segments might have a time stamp and a flag to indicate whether or not the subject was talking. Then a rule might be: If (time > 11:00) and (voice on) then display the segment.



FIGURE 3. Live Video Is Captured from a Video Camera and Presented on the Researcher's Screen

MULTIMEDIA COMMUNICATION

Electronic mail and telephones provide two separate forms of long-distance communication, both involving the exchange of information. Early attempts to incorporate video into long distance communication were based on a model of the telephone, simply adding video to the audio channel. A different strategy is to incorporate video into on-line message systems, such as electronic mail, on-line consulting systems, and bulletin boards. The latter applications enable users to exchange information asynchronously, rather than setting up a synchronous link. If video is included, the video annotation techniques described earlier can be used to help recipients identify interesting messages and handle their mail more efficiently.

Scientists are likely candidates for using multimedia communication systems, because they face an increasing need to compare their data from photographs and video sources with that obtained by other scientists. The Neuroanatomy Research Database, developed by Steven Wertheim under a faculty grant from Project Athena, provides a set of multimedia messages and a set of communication requirements that can be used to define the functionality of a multimedia message system.

The Neuroanatomy Research Database

The Neuroanatomy Research Database began with a database of terms about the anatomy of the brain, including over 1200 text definitions of anatomical parts, video of a brain dissection, slides of brain cross sections, and a three-dimensional model of the brain reconstructed from these brain sections (Figure 4). Video portions of the database could be located remotely and accessed by any video workstation through the Galatea video network server. A single image or video sequence can be incorporated into as many electronic pages as appropriate, without duplicating the images. Future work will extend the database to enable scientists to include images from their own experiments and compare them with existing images.

Pygmalion: A Multimedia Message System

Pygmalion is a multimedia message system based on a three-dimensional model for on-line communication systems developed by Wendy Mackay and Win Treese [7]. The model allows users to control message exchanges along three dimensions: synchronous or asynchronous, stream or database and control by sender or the receiver. All messages, including text, graphics and video are stored in a database at a central post office and a pointer to the message is sent to the user. The user has the impression that individual messages are local, but they are actually stored whether it is most convenient in the network. Current technology requires the video network to be separate from the computer network, but conceptually, they can be treated in the same way.

The first test of Pygmalion will be at the SIGGRAPH

'89 Conference, where a preliminary version will be made available to conference attendees. Over 30,000 people will be able to send and receive multimedia messages from a distributed network of workstations. The software is being developed by individuals from MIT's Project Athena, the Media Lab, MIT's Sloan School of Management, and Digital's Cambridge Research Laboratory.

An important aspect of managing information overload at the conference is to provide a balance between control over access to messages between senders and receivers. Senders who are willing to individually identify the recipients of a message can send messages directly, as in electronic mail. Senders who want to reach a large audience must annotate their messages and send them to a central database, as in an electronic bulletin board or on-line conferencing system. Recipients then receive personally addressed messages and can browse the database for messages of interest. Attendees will be able to record and edit messages containing video as well as text and graphics. An important feature is that all the messages can be edited by both the sender and the receiver of the message, which allows for shared control of the information.

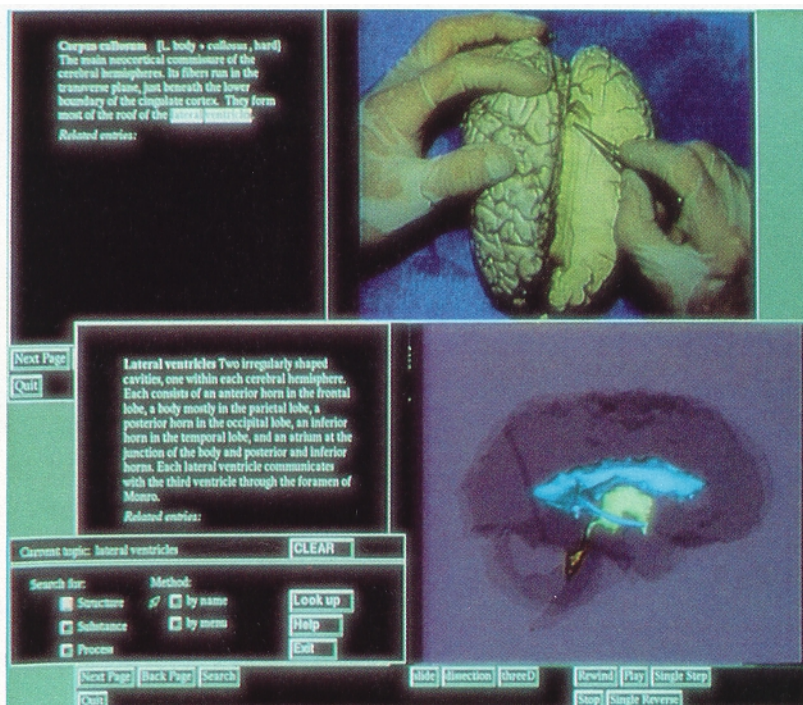
Future Directions

A multimedia message system designed for a one-week event such as SIGGRAPH '89, must address different design constraints than a system designed for daily use. However, since large scale implementations of the latter are not currently technically feasible, implementation of the former may provide useful information about the design considerations for such systems. Pygmalion will allow us to explore balance of control issues between senders and receivers of messages, the effectiveness of different annotation and retrieval schemes, particularly for multimedia messages, and the management of scarce video resources.

Video takes a lot of room. A laser videodisc, which holds 54,000 frames, has only 30 minutes of video per side. Write-once disks are available, but the per-disk cost is high. The speed of access to the video is important, and has implications for how the video is laid down on the disk. The Pygmalion system stores only one copy of each message and mails a pointer, rather than the actual video, text, and graphics. (Note that this model works at Athena, which has 850 workstations in a distributed network, but does not address the problems of sending multi-media messages over larger distances.)

Sharing presents a problem in that one person's annotation may not make sense to another person. Also, message senders must perform extra work to annotate messages, which decreases the likelihood that it will be done. Message templates, with fields to be filled in, are one approach. Another is to provide the option of asking the sender for annotations before the message is sent. What if two people try to manipulate video at the same time? It depends, of course, on whether they're

FIGURE 4. Accessing One or More Electronic Pages that Contain Text, Graphics, and/or Video



cooperating or competing, and what they're trying to accomplish. It also depends on whether they are making decisions about a fixed order of video segments, e.g., looking at different views of the same visual database or providing different kinds of annotations. In all of these cases, the participants are working together to construct a shared information environment.

CONCLUSIONS

Computer control of analog video has been commercially available to broadcasters for almost two decades and in personal computing environments for almost a decade. The requirements for interactivity and flexible manipulation of video and associated data are increasing and emphasize the limits of current technology. Analog video storage on optical videodiscs is expensive and videotape does not provide sufficiently accurate or precise control. Cooperative work applications, as in shared authorship of multimedia documents, and video prototyping, as in interactive design of multimedia software, are promising new application areas that will extend the requirements of video as an information medium.

New techniques for compressing video and encoding content and other information into the signal promise to both decrease the costs and increase the possibilities for future applications. It is important that assumptions about current and future video-based applications be taken into account when choosing among compression tradeoffs. Some applications require real-time decompression and can afford to wait for a long compression process. Others may require the reverse. Different

algorithms cause delays at different points in the compression/decompression cycle. Thus, some applications that require high quality images and long video segments may not find a 30 second preliminary delay problematic. Other applications may require very rapid changes among short segments, in which delays of several seconds may not be acceptable. The importance of image quality varies across applications, and the advent of HDTV may further change the requirements. All new technologies present tradeoffs between costs and features. Before deciding on those tradeoffs, it is important for manufacturers of digital video equipment to understand the ways in which it will be used. The tools and applications presented in this article are biased toward a constructivist point of view, which maximizes user control over the media and the environment. Digital video offers the potential for significantly enhancing them all and expanding our views of interactivity.

Acknowledgments. The authors would like to thank Geoffrey Bock, Andy Lippman, Lester Ludwig, and Deborah Tatar for early comments on this article. The work described here occurred in a stimulating environment at MIT and has been influenced by many creative people. We would like to thank Hal Birkeland, Hans Peter Brondmo, Jeff Johnson, Patrick Purcell, and Ben Rubin at the Media Lab; Ben Davis, Matt Hodges, Evelyn Schlüsselberg, Win Treese, and Steve Wertheim from Project Athena; and Mark Ackerman, Dan Applebaum, Brian Gardner, Brian Michon, and Russ Sasnett who have been associated with both the Media Lab and Project Athena.

The authors would also like to thank the participants and reviewers for the ACM/SIGCHI Workshop on Video as a Research and Design Tool, particularly Austin Henderson, Deborah Tatar, Raymonde Guindon, Marilyn Mantei, and Lucy Suchman.

REFERENCES

1. Backer, D., and Lippman, A. Future Interactive Graphics: Personal Video. Presented to NCGA Conference, Architecture Machine Group, MIT, Baltimore, Md., June, 1981.
2. Davenport, G. New Orleans in Transition, 1983-1986: The Interactive Delivery of a Cinematic Case Study. International Congress for Design and Planning Theory, Film/Video Group, MIT Media Laboratory, August, 1987.
3. Hodges, M. *The Visual Database Project: Navigation*. Digital Equipment Corporation, Bedford, Mass., May, 1986.
4. Lippman, A. Movie Maps: An application of the optical videodisc to computer graphics. SIGGRAPH '80 Conference Proceedings, Architecture Machine Group, MIT, July, 1980.
5. Ludwig, L., and Durin, D. Laboratory for Emulation and Study of Integrated and Coordinated Media Communication. SIGCOM '87 Conference Proceedings, SIGCOM, Stowe, Vt., 1987.
6. Mackay, W.E. Tutoring, Information Databases and Iterative Design. In *Instructional Designs for Microcomputer Courseware*, D.H. Jonassen, Ed. L. Erlbaum Associates, Inc., Hillsdale, N.J., 1988.
7. Mackay, W.E., Treese, W., Applebaum, D., Gardner, B., Michon, B., Schlusberg, E., Ackerman, M., and Davis, D. Pygmalion: An Experiment in Multi-Media Communication. SIGGRAPH '89 Panel Proceedings, Boston, MA, July, 1989. Special event to be presented at SIGGRAPH '89.
8. Mackay, W.E. EVA: An Experimental Video Annotator for Symbolic Analysis of Video Data. *SIGCHI Bulletin* 21, 1 (Oct. 1989). To be published in the Special Issue on Video as a Research and Design Tool.
9. Rubin, B., and Davenport, G. Structured Content Modelling for Cinematic Information. *SIGCHI Bulletin* 21, 1 (Oct. 1989). To be published in the Special Issue on Video as a Research and Design Tool.
10. Sasnett, R. Reconfigurable Video. Master's thesis, Department of Architecture, Massachusetts Institute of Technology, 1986.
11. Mackay, W.E., Ed. *Workshop on Video as a Research and Design Tool*. ACM/SIGCHI, 1989. In press.

CR Categories and Subject Descriptors: D.2 [Software]: Software Engineering; D.2.2 [Software]: Tools and Techniques—*user interfaces*; D.2.6 [Software]: Programming Environments—*interactive*; E.m [Data]: Miscellaneous—*video data*

General Terms: Design, Human Factors

Additional Key Words and Phrases: Digital video, video

ABOUT THE AUTHORS:

WENDY E. MACKAY has worked for Digital Equipment Corporation for 10 years. Currently, she is a doctoral candidate at MIT under a scholarship from Digital. Her interests include multimedia education and communication. Author's Present Address: MIT Project Athena, E40-366, 1 Amherst Street, Cambridge, MA 02139.

GLORIANNA DAVENPORT is an assistant professor of media technology and director of film/video research at MIT's Media Laboratory. She has produced, shot, and edited a number of observational documentary movies and has been involved in the development of interactive editing and delivery systems since 1983. Author's Present Address: MIT Film/Video Group, The Media Laboratory, E15-432, 20 Ames St., Cambridge, MA 02139.

Permission to copy without fee all or part of this material is granted provided that the copies are not made or distributed for direct commercial advantage, the ACM copyright notice and the title of the publication and its date appear, and notice is given that copying is by permission of the Association for Computing Machinery. To copy otherwise, or to republish, requires a fee and/or specific permission.

ACM VIDEOTAPE ON INTERACTIVE DIGITAL VIDEO

Announcing

Announcing

ACM is developing a videotape that will further demonstrate the *technologies discussed in these pages*. In addition to serving as a *video supplement to this issue*, the videotape will contain the following material:

- Excerpts from **Palenque** (*an Interactive History Lesson*)
- The **Carnegie Mellon Software Engineering Training Program** using interactive video
- Samples of the state-of-the-art work being done at the **MIT Media Lab**
- The use of video compression and playback developed by **Intel**

Circle #118 on Reader Service Card