INTERACTIVE MULTIMEDIA ON A SINGLE SCREEN DISPLAY

Glorianna Davenport,
Film/Video Group of the Media Laboratory
Massachusetts Institute of Technology

For: Videotechnology Technical Session: *Current Applications of Videotechnology in Computer Graphics Systems,* National Computer Graphics Association Conference, March 22, 1988.

ABSTRACT

Interactive delivery of multimedia material destined for educational courseware or large reference archives imposes complex constraints on both the delivery system and on the application design. *A City in Transition: New Orleans,* 1983-86, a cinematic case study of urban change, combines 3 hours of movie sequences; a still frame library of characters, places and maps, mastered on optical videodisc; a wide variety of text documentation; and relevant demographic and economic statistics.

Intended for interdisciplinary study of political, socio-economic, and design issues, the project requires that students be able both to use available material and to add new visual material to the database. The interactive version is currently being developed at the Media Laboratory on a Project Athena Visual Workstation, runs under X-windows with a Parallax 1280 board a single-screen high-resolution display. This paper will describe issues and solutions related to clarity and directness of information access; time-linked relationships between movie segments, graphics and text; and the set of user tools which allow viewers to edit video footage, create graphic models and link their notations to the master database.

INTRODUCTION

As a moviemaker and movie editor, the idea of an interactive, or interruptable, movie environment has methodological association both in 1) shooting and 2) editing. In the first case, we record a bit, wait, record a second bit; in the second case, we take all those bits, shuffle them in various ways and assemble them with an eye to creating an experience for our viewers.

Interactive environments for viewing

One criterion for evaluating a movie involves the fascination and intensity of thought the movie generates for a particular audient. Frequently this initial experience is enough; however, in the case of certain documentary movies, a particular audient will want to know more, will want to really study the situation which has been commemorated on film

or tape. In their traditional linear form, we lack both the additional knowledge and the means of precise review which allow us to study a given reality in depth.

The information experiments designed by Vannevar Bush, Doug Engelbart, Theodore Nelson and others over the last quarter century contend that parallelism in thought could/should find its analog in access to digitally stored information. Today, in an overwhelming tribute to these visionaries, we are deluged with conferences devoted to hypertext and hypermedia. The fact that these ideas are being embraced so intensely owes as much to the distance we have traveled in hardware design and system software as it does to our acculturation into the information era. Not only can multiple text and graphics sources be displayed concurrently, but also motion picture and sound have been added to the information palette. Anew age has dawned. Ostensibly we should be able to don our creative beret, academic robe, or Sherlock Holmes hat and pipe and participate in our own style of the information fare.

Ivan Illich in H2O and the Waters of Forgetfulness writes:

Each culture shapes its own space, the very space it engenders in becoming a culture. Space is a social creation which results from the all-embracing asymmetrical complementarity enshrined in each culture.[1]

Clearly one of the central issues for today's information culture has to do with defining a new kind of space --a viewing space for complex, multi-layered information sets; a thinking space which will help us model meaning; a space which supports both the "hands on" and "heads in" activity of learning.[2]

Time-dependent media (sound and picture) add enormous vitality to this space but they also add complexity both in terms of hardware and systems construction and in terms of designing interfaces which support personalized perception. Today as we shape new media spaces we can begin to reflect on what the medium itself means. To that end I will offer some observations about the moviemaking process and interactive media, and discuss some aspects of *A City in Transition: New Orleans,* 1983-86, a small multimedia resource based on a cinematic case study which we hope will have some bearing on tomorrow's feast.

BACKGROUND: The Movie-Maker's Methodology

As an observer, filmmaker or viewer, I am curious to learn why and how a given individual does what s/he does, and how the cumulative effect of multiple interactions changes the nature of our environment. However, given the tradition limits of linear

---

[1] Ivan Illich, H O and the Waters of Forgetfulness, Heyday Books, 1985.

[2] Seymour Papert, Mindstorms, Basic Books, 1980.

viewing, the two observational roles -- that of moviemaker and movie viewer -- are fundamentally different.

The observational filmmaker's working method involves researching the who, where, when, and what of a potential subject. As the filmmaker records "aspects of what takes place in the presence of a camera",[1] s/he hopefully gains a deeper understanding of what and begins to formulate why. I will call our original recorded observations our first level interpretation. A second level interpretation is introduced when we edit our observations into a form, a story. Traditionally the physical linearity of this second level interpretation has defined an inviolate temporal experience which generates the third level interpretation of any viewer. As viewers, we watch, experience and think about the movie but are unable to take an active role in shaping our discovery.

Random Access and Movie Editing

The incorporation of "movies" into interactive environments began with the introduction in 1979 of the CA V (Constant Angular Velocity) optical videodiscs and players which could be serially controlled. Early genres in the medium included surrogate travel, the "electronic book", and corporate training movies. However, for moviemakers one of the most interesting potential applications for the medium was video editing.

The ability of CAV optical videodisc players to rapidly search to frame-accurate locations (2 1/2 seconds or less per 30 minutes of material) make them ideal video storage devices for virtual editing systems, if we have several players on line. By virtual we mean that we can view an edit-list seamlessly before recording it, thereby avoiding the cumbersome and creatively draining cuing times required for videotape editing. Using an associative or relational database we can also name shots. In more sophisticated environments, we can grab specific video frames, shrink them to postage stamp size, and use them as digital icons to represent either specific shots or characters and places. Our research over the past three years has shown this to be a particularly effective way of personalizing a user's access to video material so long as the icons remain unambiguous to the viewer.

Random Access and Movie Viewing

The incorporation of video into integrated multimedia environments requires a radical reinterpretation of the idea of authorship in motion pictures. Although home VCRs encourage viewers to record programs and review them, they are not designed for fine editing. Computer environments which support multiple video playback devices (with the fastest frame-accurate seek times possible) are, given the right control software and user interface, the equivalent of professional video editing systems with the added advantage of allowing virtual viewing. When we add a database referencing shot location, content, and context, as well as ancillary text and graphic materials which clarify the narrative, we create a hybrid media retrieval and presentation which allows viewers to reconfigure information at will.

[1] Richard Leacock, in conversation.

Moviemakers who work hard to create linear stories designed to entertain and envelop generally find the notion of giving up control distasteful, if not shocking. As we look towards a future of enabling viewers to access what they want, when they want and how they want, I support the efforts of my colleagues. The creation of beautifully framed images, evocative sound tracks, and coherent stories (edited movies) are personal interpretations which transmit memorable experiences; they are a desirable and needed resource. Just as not everyone can or wishes to go everywhere with a camera, not everyone wishes to spend their time editing or searching for specific information. It is my contention that, even in interactive environments, it is valuable to publish a formatted set of observations. Those who wish to share the holistic experience will. At this level, the abyss associated with "random access" and virtual editing is obscured.

Given the future potential of random retrieval environments, I propose that the bogy is less that users may alter the filmmaker's constructed reality than it is discovering what kinds of movies really benefit from the effort involved in databasing and "on line" collections of supplemental documentation.

Partially in an attempt to explore this bogy, I embarked five years ago on a three year cinematic case study, "A City in Transition: New Orleans, 1983-1986." Shot by Richard Leacock, the project explores the process and issues related to the development of the downtown riverfront. Today we are implementing a 3-hour interactive version of this case study.

APPROACHING THE CASE STUDY: CINEMATIC CONCERNS

Why a new project? Would not a pre-existing program have sufficed? Perhaps, but two aspects of using a preexisting program were troublesome. Firstly, one could not study the process of the filmmaker's research methodology first hand. Secondly, I wanted to use as the basis of the experiment a subject which would have hooks into as many aspects of our cultural history as possible. I wanted to make a "thick description" which could continue to grow long after the active process of cinematic observation had ceased. By focusing on urban change I hoped to make a visual database which was relevant to the fields of architecture, planning, economics, political science, sociology, and human behavior. Serendipitously, MIT's School of Architecture and Planning would offer us an active and immediate audience.

New Orleans at that time offered some of the best possible conditions for observational filmmakers. The focus could be contained geographically (downtown riverfront), and temporally (it was 1982 and a World Exhibition was scheduled to open in 1984; this was catalyzing development activity). The population was small (600,000) and many of the major players readily agreed to provide the filmmakers access. The culturally rich heritage of the downtown area juxtaposed to the fragile current economy of this mythic city suggested that our story might invite future research.

Although my initial concept has matured considerably from the technical perspective, I correctly assumed from the start that the cinematic portions of the project would for the

most part reveal moments in which people were negotiating, and that many of these moments would involve public gatherings. While cinema is great for communicating personality and style, it is not well suited to factual commentary or second hand interpretive analysis. For serious study, most of the scenes we were recording would benefit by being associated with a variety of support information, such as maps, charts, spread-sheets, zoning documents, newspaper articles, and the like. Many of these we collected as we were filming; they provided us with essential information about our story .The suggestion that real-time video segments were appropriate to undergraduate design education was embraced by Joseph Ferreira who was founding the Computer Resource Laboratory in the School of Architecture and Planning with the support of Project Athena. This recognition was essential as I endeavored to develop the computational version of cinematic case study as an educational resource.

THE INTERACTIVE ENVIRONMENT

The challenge of designing an information space which supports cinematic integrity while allowing users the option of stepping through the information, of asking questions, reconfiguring sequences, annotating scenes and aesthetically integrating this information back into their own observations raises some interesting questions about how and what the cinematic experience is about.

In 1983, as Ricky Leacock and I began our case study, Digital Equipment Corp. and IBM funded "Project Athena." This 5-year program was committed to the development of a coherent, networked system for undergraduate education. The possibility of running movies in a computational environment was exciting; it meant that we could design movies for a new audience. In the early days of Project Athena, we talked a lot about our ideas and heard a lot about what the ultimate environment would be like; however, in 1984 our only research machine was a IBM XT, later upgraded to an AT. Our first task was to design a videodisc controller and a screen interface. As we began to play with discs which we made or were given (such as the NASA disc), we discovered the power of using a database manager to make edit logs and edit decision lists;[1] we appropriated dBase III while we awaited Project Athena's selection of a database of choice.

In the spring of '85 (as I recall) Project Athena announced its decision to support Ingres on the network, and later the same year the Visual Courseware Group became an official Project Athena entity. The charter of this group, which was given its own budget, staff and facilities, was to create the technical specifications and write general code which would insure a standard video environment for the handful of courseware projects which required video sources. Under the guidance of Rus Gant, hardware specifications were drawn up which integrated color graphics and video on a single screen display into the chosen Athena environment of Unix and X-windows via the Parallax 1280 board. To Project Athena, the single-screen display was an essential feature of any workstation. In hindsight, I find it somewhat ironic that I was ambivalent to the concept of a single screen display. The potential clutter of viewing and editing a movie on the same screen and simultaneously writing a paper or searching for a text document terrified me. In fact,

[1] Russell Mayo Sasnett, Reconfigurable Video, MSVS Thesis, Massachusetts Institute of Technology, 1986.

the only advantage I could envision clearly was the potential for overlaying a 3D graphics on the video image and the ability to quickly grab and view video icons.

Today, having worked in the single screen delivery mode for nine months, I am excited by the flexibility of the X window environment. As we develop tools which support our theories about learning environments and movies, we are simultaneously observing students as they acquire knowledge using our environment; they in turn reflect their experiences back into the design framework.

CASE-STUDY APPLICATIONS IN AN ACADEMIC CONTEXT

Before pursuing the interface goals, it is helpful to look at the goals of a case-study approach to curriculum. In general, the case-study method suggests that students benefit from the study, analysis and problem solving presented in real-life situations. For example, the method is commonly used in the study of law and management; students are given sufficient information on a particular situation, are asked to evolve a strategy, and can compare their solutions to what actually happened.

The New Orleans project was used in a similar way in an Introduction to Urban Planning course last term. Taught by MIT professor Dennis Frenchman, the course is designed to allow students to develop a methodology for looking at cities, assessing development issues, and offering the city philosophical and design alternatives for development. As future planners, students need to be able to produce classical analyses of urban spaces which reflect and illustrate their proposed design aesthetic. In his opening lectures, Dennis presented students with historical patterns of urban development, and a set of research methodologies with suggested information sources (maps, media, interviews, official reports) which could be used to learn about a city: how it had grown, how it functioned, what citizens liked or disliked about it. His first assignment required that students research the geographic, historical, economic and demographic profile of a place; in the second and third assignments students were required to offer a design concept and suggest a practical approach for selling the concept to the various regulatory agencies and the citizens.

At the time of the second assignment our movie had just been mastered on videodisc and the available electronic implementation was miniscule. Those students who choose to work with us were brave souls. The video movie ran in a video window on the Parallax 1280 video-graphics board. Our video editor was, and still is, somewhat cumbersome, particularly in terms of access to the button speed controls; however it does provide all the basic facilities of a video editor including defining an event, adding an event to a list etc. Students used emacs for taking notes and writing their paper. However, almost all the available support documentation was stuffed into a cardboard box and our Ingres database was not even on-line for video searches. It was a far cry from our dream.

Last term one presentation was completed on line. When I watched it I was stunned, as was Professor Frenchman. The alternation of critical analysis and the immediacy of being there and "knowing" the characters was overwhelming, particularly given the rigidity of the presentation environment, which caused some problems. The students used a button

box with successive buttons for each video segment included in the text. Only one disc player was on-line, which meant not only that one could not view an edit-list seamlessly, but worse, as the New Orleans project is contained on six discs, discs had to be changed frequently during the presentation session. Removing the active video signal from the system causes other chaos, which I will not go into as we are currently putting 8 videodisc players on-line, which will obviate all of these problems.

Despite obvious flaws in "transparency," both Frenchman and the students were thrilled by their experience. The students learned "how important the personal is in creating change" which after all is shaped as much by particular personalities as it is by the power of the theoretical goals.[1] Student users of the development system provided an invaluable perspective about front-end tools for data management; in particular, I came to realize that viewers of presentations designed for the workstation also need access to the referential information and an interface for accessing this information should be generic to all levels of use.

DESIGN CHOICES / SCREEN LAYOUT

Let us jump into our most recent version of user mode. What you will see is a simulation of icon and text selections which have been generated by simple queries: who, what, where, when. As a student we begin by viewing a short authored scenario which can be found under "assignments." The student's job in this case is to research all relevant aspects of this urban situation including the site, the individuals and agencies involved in the decision making process, the procedures and precedents, the goals and dreams. What follows is a short video, which will hopefully demonstrate one way in which viewers can travel through the information space. At this point we are still running with one player; the search times will soon become invisible.

The date is January 1983 *and you have been hired by a group of concerned citizens to advise them on the impact that a proposed development will have on their neighborhood.*

What you see is the text and the indication that a video segment has been linked to the assignment by the professor. We press the available button to see the video footnote. Now you become the pilot.[2]

In this version, we have arbitrarily chosen a screen layout, which we feel, is appropriate for novice users of the system. The left-hand side of the screen is fixed. Video runs in a window in the left-top quadrant of the screen; below it we have a controller which is not designed for editing but rather for browsing, and text which is related in the database either to a specific icon or to the video sequence. We have reserved the right-hand side of the screen for query modules. At this time, we are using a limited number of fixed and dynamic video icons, which are represent aspects of the scene or story.

---

[1] Don Knerr and Jon Teischer, Urban Process: The Jax Brewery Conversion, New Orleans, Louisiana, for 4.7361 (Frenchman/Dash), January 28, 1988, MIT.

[2] Benjamin Rubin (MIT, MS '89) and Brian Bradley (MIT '89) deserve the lion's share of credit for this design. Brian Bradley has also been largely responsible for the design of the remote data structure.

Icons can be used to specify particular searches; for instance, mousing portrait icons yield resumes and short video clips of the person. Although icons can only be relied on for simple searches where their meaning will not be ambiguous, they do represent some data categories, such as characters, very well. In the next software version users will be able to choose between icons and text in the form of pull down menus. In combination, these will create complex searches: for instance a person in a particular role or at a particular event, or all on line scenes of the person. The icon window will be overridden by pull-down data-structure windows, or whenever a viewer needs access to another window, a text editor for note taking, a window for transcript follow, or a video editor which allows users to create lists.

Of course, more expert users can change the screen layout.

DATABASE SEARCH

As I mentioned earlier, the central ideological and practical tenet of research involves the use of a relational database. Philosophically and practically, if movies are to be relevant to a wide variety of researchers, the database must allow for expansion of the information set, in cases such as: new information related to downtown development in New Orleans; other visual databases including on-line slide libraries; expert opinion; personal notes. Finally an extensive database of shot characteristics is essential to the future existence of an automatic media-maker.

Designing the database structures constitutes the most difficult task conceptual task; this is followed by the manually arduous task of data-entry. Currently we are in phase II of designing/testing data-structures for content. Current data frames focus on: people (individuals, groups, rolls); places (buildings, streets, zones, rolls); actions (political, legal, physical); issues; events; and references. Pull-down menus, Macintosh style, were chosen for the data interface because they clearly describe the hierarchical nature of the data in each category. Full query implementation will, of course, require dictionary and thesaurus modules.

Technically, on the Project Athena system, data in Ingres is on a remote main frame. We use a program called GDB written by Noah Mendelson to access Ingres from our local machine. The main advantage of using GDB is that multiple searches can be progress concurrently, and personal database entries can be joined to the master database. Disadvantages include slow seek time and network crashes. We have discovered that the disadvantages will make it necessary for us to appropriate some database functions to occur locally. Currently, local functions include icon lists, edit-lists, and flat files of support material. We will shortly implement scene and character tables locally.

For linked time-dependent displays, such as transfer follow, several approaches may be taken. One solution is to use an authoring system such as MUSE, which Matt Hodges is currently developing at Project Athena. Muse is structured to allow users to place several levels of event display into a time line. If this program is flexible enough we may also be able to link it to database returns through specific rule-based operators. This might be an interesting way for the machine to follow and react to particular passes through the information space.

CONCLUSION

The New Orleans in Transition project, although still in a rather crude state, presupposes that active and curious inquiry is an important feature of any learning environment. Over the next year we will be able to study how students respond to this environment, and learn how to make the system more responsive to personal styles. Already our students are implementing a new generation of video editing interfaces: using an icon palette, users will be able to script scenarios and view these scripts in time lines. Other students are interested in integrating additional personally generated visual input sources. Other students are focusing on data returns and their representation. In watching the workstation environment evolve, I can, as a moviemaker, finally reflect on what "random access" means to shaping movie forms for the future. And although each stage of the design formulation is exciting, I also can now look forward to finding new stories to record.