COMPUTATIONAL VIDEO: TELEVISION AS A HUMAN INTERFACE*

A. Lippman

Massachusetts Institute of Technology

ABSTRACT

The advent of the optical videodisc, and the simultaneous development of man-machine communicative devices now provides the opportunity for the development of a new medium unprecedented in its descriptive ability, and unlimited in its scope. As with any medium, the requirements for authoring, the rules for use, and the applications are areas for research. This medium is termed "Computational Video", and it envisions the use of the video display as a complete user interface, and the processing behind it as video-domain image and data system.

Two examples of work in this field are described: the Movie-Map and the Movie-Manual. The first is an early test of making movies, edited on-the-fly to familiarize a person with a new topographical environment. Presentation is by real imagery, computer generated animation, and processed pictures, all sound synchronized. The <u>experience</u> of a place rather than its schematization is primary.

The second is the application of this medium to instruction and training. In this case, the disc-computer-display medium performs as tutor and assistant, creating continuous, interactive descriptions of maintenance and repair tasks. The manual is responsive, predicating its presentation on a dynamic model of the user and sensing events in the real world. The optical videodisc is used as a storage device for pictures, sound and data, and all modalities are symmetrically available for presentation and input.

RÉSUMÉ

L'avènement du vidéodisque optique et le développement simultané des dispositifs de communication homme-ordinateur ont permis la mise au point d'un nouveau support d'information sans précédent quant à sa capacité descriptive et de portée illimitée. De même que pour tous les supports d'information, les besoins de création, les règles d'utilisation et les modes d'application sont des secteurs de recherche. Le présent support d'information est appelé "vidéo de calcul" et il permet d'envisager l'utilisation de l'affichage vidéo comme une interface complète pour l'usager, et le traitement qui l'entoure comme un système d'images vidéo et de données.

Voici la description de deux exemples de travaux dans ce domaine: La ciné-carte et le ciné-manuel. Dans le premier cas, il s'agit d'un des premiers essais de cinématographie préparé au pied levé pour familiariser un utilisateur avec un nouvel environnement topographique. La présentation est faite par imagerie réelle, par animation informatisée et par images traitées par ordinateur, le tout étant synchronisé sur le plan sonore. <u>L'expérience</u> d'un endroit plutôt que sa schématisation est essentielle. Le deuxième exemple consiste en l'application de ce support au domaine de l'enseignement et de la formation. Dans ce cas, le support affichage-disque agit comme un précepteur et son assistant en créant des descriptions continues et interactives de tâches d'entretien et de réparation. Le manuel donne des réponses en basant sa présentation sur un modèle dynamique de l'usager et en détectant les événements dans le monde réel. Le vidéodisque optique est utilisé comme dispositif de stockage pour les images, le son et les données et toutes les modalités sont symétriquement accessibles pour la présentation et l'entrée.

* *Work reported herein is supported by the Defence Advanced Research Projects Agency Contract No. MDA903-78-C-0039, and the Office of Naval Research Contract No. N00014-75-C-0460.*

Computational Video: Television as a Human Interface

1.0 Introduction

Desired is the creation of a "window" into a computational space. The purpose of the window is to allow the viewer to usefully peer into information in a way heretofore closed to him. Through this window is a "dataland" replete with images, knowledge and utility. It has been closed because the past builders of such windows have been computer programmers, and the predominant users of such information systems have been other computer programmers, or operators; never simply people.

It is now possible to change this -- and this paper will elaborate on several alternatives. The dark world of computerized information may now be accessible. The means exist, and the direction of these means is developing. The specific approach to be described uses familiar elements -- a television screen, for example -- and familiar controls -- touch, speech and gesture. The computer enters as an invisible mediator, a personalized guide and tutor; the medium is defined as the television transmission path. But first, the background for this change.

1.1 Unifying Trends

Computer graphics, television, image processing, even publishing, once widely divergent studies and disciplines are gradually merging. An agency of this merger is the recognition by the computer graphics industry of the need for and the virtue of raster scan display, coupled with technological capability. Full color video displays, being linked to ever developing semiconductor technology, are costing less and doing more every day. Further, display techniques are improving to the extent that high enough information content in a raster scan display and the ability to portray solid objects is winning over the engineering design community, long a stronghold for calligraphics.

A second force in this direction is the growing use (at least industrially) of the optical videodisc. Admittedly, this is not quite the avalanche one would wish for: a great percentage of industrial disc players "in the field" are in General Motors showrooms, but it is a beginning. A growing part of the market consists of small scale users who are experimenting and gaining necessary expertise.

Users of discs fall into two readily distinguishable categories: those who view the device as an extension of available audio-visual delivery media, and those who envision it as a mechanism capable of wholly new interactive scenaria. The possibilities of the second approach will be elaborated upon later in this paper, but both are a significant, even driving force.

A final component is the increasing incursion of computers and computing into our daily lives. It is only necessary to cite the growing popularity of home computers, paltry and deficient as they are, to prove this point. There is, however, more compelling evidence. Computers are becoming demystified. No longer are they the center of attention of systems in which they participate. They are becoming components, visible by their effects rather than their stated presence. As examples: the modern microwave oven is not a computer that cooks, but an oven that thinks. Similiarly, an automobile is not a rolling computer center, but a car, with computer controlled carburetors, speedometers and sound systems.

1.2 Component Definition

If the previous three observations are accepted, then the definition of the computational video interface is almost complete; the ingredients are acceptable and available. They are (1) a computing system acting as invisible mediator; (2) a use or appearance which the system can assume; (3) the television system as the user interface. Necessary, in addition, is a more precise definition of what is meant by television, how it is used, and what it can be.

Firstly, television, as a display, means broadcast compatible, standard resolution video. The confluence of raster scan computer graphics and the optical videodisc demand this: the disc is a broadcast compatible device. Past this mandate, however, is a sufficiency. It is becoming increasingly clear in the graphics community that for a great many applications, standard television may well be good enough [1,2]. It is clear that the cost of using a higher resolution standard does not result in a comparably clearer, sharper, or denser display. Thousand line television does not display twice as much as 525. It does, however, cost more than twice as much in computation, storage, and display complexity. It also eliminates the home consumer.

Secondly, television is more than a simple display prerogative. The video signal is capable of containing more than the brightness and chromaticity of a physical object. It is already the transmission path for data in home information systems, and as a captioning mechanism. The characteristics of the signal itself make it amenable for use as a high bandwidth data path and high density storage format.

When regarded this way, the argument that the videodisc, for example, provides visual richness in an otherwise graphically sparse environment may be extended to mean providing computational and storage complexity in a similarly computationally cost effective manner. This particular extrapolation will be developed through example, but suffice it to say that what is meant by television is not merely an image display format or picture coding scheme. It is a complete visual and computing arena encompassing high quality color imagery <u>and</u> the data density and interaction realms of large scale computer graphics establishments.

2.0 Examples

Two examples will be presented. They are both characterized by the incorporation of the optical videodisc into an integrated processing environment. In the first, the system exemplifies a prototype interactive movie. Two points are stressed. The system functions not as a computer, but as something else entirely, a map. Also, a compendium of techniques are used to create a computational equivalent to a dramatic movie presentation.

In the second, man-machine interaction results in something more than explicit control. A personalized system is created. Similar components to those used in the movie-map are again used to make a personal, electronic repair manual. The manual is analogous to a printed one: it becomes the property of a single owner and changes with time to suit the needs of that person. The changes are both explicit and implicit additions and annotations. The manual takes a presentational form dictated by the joint experience of the computer system and the manual user.

2.1 The Movie-Map

The movie-map at its simplest level is a surrogate for a printed topographic chart [3]. Its purpose is evidently the same: to familiarize a stranger with a new geographic place. The mechanism of this familiarization, however, is essentially different. Rather than present the "map-reader" with a bird's-eye view of the terrain implemented via labelled lines on paper, the reader is engaged in an interactive visual tour. He is allowed to "drive" through the space, seeing what he would see were he there, and inquiring about any feature of the environment encountered along the way in greater detail than can be accomplished by simply driving past.

As a map, the goal is to simulate the experience of a visit in so much detail that when encountering the actual place, the feeling engendered is one of having been there before. To implement this vision, a computer with access to a relatively complete database of the town directs the operation of two videodiscs used alternately, to create a continuous, contiguously variable tour. Concatenated together are individual frames of forward travel, negotiation of turns through intersections, and segments in different seasons or presentational mode. Additionally, the same discs intersperse close-up views of buildings seen *en route*, and audio-visual depictions of their interiors. A recent augmentation is the ability to integrate this tour with a dynamic topographic representation. By touch command, one may switch from the tour to a map, travel in a lattice over this map, and zoom into any intersection. The route traversed is overlaid onto aerial images and maps [4].

The point made by this system is that a continuous presentation that is apparently cinematic may, in fact, function as something else entirely. The components may as well be used to make a map or be seen as a prototype for a more general interactive movie maker. The key is that the system responds rapidly to any of a class of user interventions, made anytime. Rather than epitomize the "present-query-continue" mode, continuous involvement is encouraged, and continuous response is the result.

2.2 The Movie-Manual

The movie-manual is an extension of the concepts developed in the map to training a task. Variability of presentational mode and content exist, but they are incorporated into a system that is the tool of a single owner [5]. It becomes his personal property, and takes an appearance dictated by past interaction and modification. Unlike the map, which is the same to each user, and draws from the same material for each reference, the manual acclimatizes itself to the peculiarities of one person. Explicit verbal and graphical annotations are digitally stored, and reappear, in place, on the next perusal. Descriptive modes are chosen based upon past experience.

Also, the manual is a more general composition system. Its predicates for movie accumulation are based not upon a global geographic reality, but upon the individual element's conceptual and independent rules. The ability to indicate articles for elaboration, for example, depends upon the contents of each possibly unrelated frame. Any unfamiliar mechanical assembly or tool showed in a sequence is fair game for inquiry. This system adds the ability to write to the hardware of the general movie-maker. The computer becomes more than a director and mediator. It also accepts changes and remembers events across sessions. The element of personalization is introduced.

3.0 The Addition of Data

Both examples presented above are characterized by the inclusion of a robust data processing component. One recreates a visit;the other provides instruction. In each case, the main interaction surface is the television screen, and the optical videodisc augments computational reality with visual and aural richness. In some sense, each may be regarded as a system inclusive of realistic imagery in an environment not otherwise capable of doing so. The map would require a flight simulator; and the manual, an encyclopedia. Yet each is realizable in a comparatively sparse graphical system by the combination of videodisc and computer.

This may be developed one step further when the disc video signal is used simultaneously to carry image and digital data.

The use of the disc to substitute for an unrealizably expensive image generator now is extended to become a substitute for a large scale computing resource. Both examples are characterized by the fact that the image currently on view dictates what images may be presented next. When this situation occurs, the video signal itself can carry this data. A portion of the image frame is used to code the database and program segments that control the play of that frame. That segment is intersected with current user inputs, and the intersection is a command describing what to do next.

In the movie-map, for example, the next turn sequence and the coordinates for each "touchable" object on view can be stored in a few TV lines, and be instantly available to process a user request. In the manual, editing controls and elaboration data may be recovered. As stated, the technical issues of incorporating this data into the video frame are not the critical issue. Captioning systems already do so. The research topic is how best to distribute the control program and database throughout the presentation. This is the completion of the notion of computational video.

4.0 Summary

The case has been made for video as a conjoint image and control medium, and the standard television as the primary graphic interface. Work is under way to implement these goals. The movie-map database has been translated into video form and is being simulated in use. Optical videodiscs have been made that encode the data into the top and bottom visible lines of the picture. The result of this work is to make a simple, extensible version of the map.

References:

[1] Andrew Lippman, "High Resolution Display on Standard Raster Scan Systems," Architecture Machine Group, M.I.T., Cambridge, Massachusetts, December 1978.

[2] F.C. Crow, "The Aliasing Problem in Computer-generated Shaded Images," Communications of the Association for Computing Machinery, 20(11):799, November 1977.

[3] Andrew Lippman, "Movie-Maps: An Application of the Optical Videodisc to Computer Graphics," Proceedings, SIGGRAPH '80, Seventh Annual Conference of Computer Graphics and Interactive Techniques, July 14-18, 1980, Seattle, Washington.

[4] Robert Mohl, "The Interactive Movie-Map: Surrogate Travel with the Aid of Dynamic Aerial Overviews," Architecture Machine Group, M.I.T., Cambridge, Massachusetts.

[5] David S. Backer, "Interactive Movies: Explicit Viewer Interaction For Dynamic Control of Movie Presentation," Architecture Machine Group, MIT, Cambridge, Massachusetts.