VIDEOPLACE--An Artificial Reality

Myron W. Krueger, Thomas Gionfriddo and Katrin Hinrichsen
Computer Science Department
University of Connecticut
Storrs, Connecticut 06268

Abstract

The human-machine interface is generalized beyond traditional control devices to permit physical participation with graphic images. The VIDEOPLACE System combines a participant's live video image with a computer graphic world. It also coordinates the behavior of graphic objects and creatures so that they appear to react to the movements of the participant's image in real-time. A prototype system has been implemented and a number of experiments with aesthetic and practical implications have been conducted.

Introduction

This paper describes a number of experiments in alternate modes of human-machine interaction. The premise is that interaction is a central, not peripheral, issue in computer science. We must explore this domain for insight as well as immediate application. It is as important to suggest new applications as it is to solve the problems associated with existing ones. Research should anticipate future practicality and not be bound by the constraints of the present.

Unlike most computer science professionals, who have been content to rely on traditional computer languages and the hundred year old keyboard as the means of input, designers of graphic systems have long recognized the importance of the human-machine interface. Even so, most innovations, including the light pen, joy stick, data tablet and track ball have been dictated by the minimum needs of immediate graphics applications.

There have been few experiments motivated by a purely intellectual desire to explore the means through which people and machines might interact, independent of specific applications. One such novel approach was Ivan Sutherland's head-mounted stereo displays which sensed the orientation of the viewer's head and displayed what would be seen in a simulated graphic environment from each position. [SUTH68]

Another unique approach was taken with the GROPE system at the University of North Carolina. It provided force feedback to a remote manipulator that could be used to pick up graphic blocks. [BATT72] In addition, there have been the well funded efforts of the Architecture Machine Group at MIT, including the Dataland project, MovieMap and "Put that there". [BOLT79, 80, 81], [LIPP80]

Finally, my work in Responsive Environments, beginning in 1969 and continuing to the present, has allowed a participant's movements around a room to be translated into actions in a projected graphic scene generated by the computer. [KRUE77, 83]

This paper describes one of my early experiments with Responsive Environments, the VIDEOPLACE project currently under development and applications planned for the near future.

Abstract Versus Concrete Intelligence

The observation underlying this research is that there are two quite different aspects of human intelligence. The first is the logical, deductive, explicitly rational process that we associate with abstract symbolic reasoning. While the technically inclined take great pride in this skill, a large fraction of the population has no interest in developing it. The second is the facility for understanding, navigating and manipulating the physical world. This ability is part of our basic human heritage.

As a greater percentage of the population becomes involved in the use of computers, it is natural to expect the manner of controlling computers to move away from the programming model and closer to the perceptual process we use to accomplish our goals in the physical world.

Early Responsive Environments

In 1969, I began to explore the idea of physical participation in a graphic world using the paradigm of a Responsive Environment. A Responsive Environment is an empty room in which a single participant's movements are perceived by the computer which responds through visual displays and electronic sound. Since 1970, video projection of computer graphic images has been used to provide the visual response.

PSYCHIC SPACE

In PSYCHIC SPACE, a Responsive Environment created in 1971, sensing of the participant's behavior was accomplished through a grid of hundreds of pressure sensors placed in the floor. As the participant walked around the room, the computer scanned the floor and detected the movement of his feet. The person's position in the room was then used to control an interaction in a graphic scene which was displayed on an 8'x10' rear-screen video projection.

In one PSYCHIC SPACE interaction, the participant's movements in the room were used to control the movements of a symbol on the video screen. After a few minutes, allocated for exploration of this phenomenon, a second symbol appeared. The participant, inevitably wondering what would happen if he walked his symbol over to the intruder's position, moved until the two symbols coincided. At that point, the second symbol disappeared and a maze appeared with the participant's symbol at the starting point. (Fig. 1)

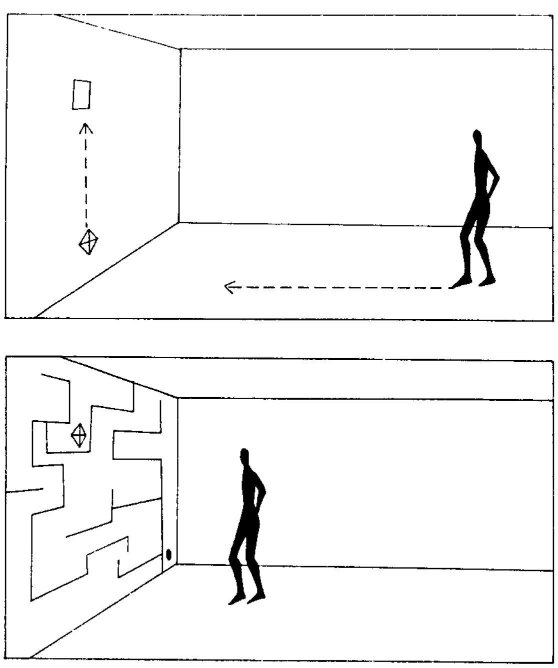

Fig. 1

Again, inevitably, the participant tried to walk the maze. However, after a few minutes the participant would realize that since there were no physical boundaries in the room, there was nothing to prevent cheating. When this realization struck, the participant, typically with some ceremony, raised his foot and planted it on the other side of one of the graphic boundaries. However, the maze program had anticipated this response and stretched that boundary elastically. Subsequent cheating attempts were greeted with a number of other gambits. The participant's symbol might fall apart; the whole maze could move; or, a specific boundary would disappear and a new one would appear elsewhere. By the end of the experience, the participant could have encountered as many as forty different variations on the maze theme.(KRUE77, 83)

PSYCHIC SPACE was presented as an aesthetic work in the Union Gallery at the University of Wisconsin. It suggested a new art form in which the participant's expectations about cause and effect could be used to create interesting and entertaining experiences, quite unlike anything that existed at that time and still different in spirit from the video games of today.

VIDEOPLACE

Concept

In 1970, I combined computer graphic images, created by an artist using a data tablet, with the live image of people. Observing their reactions to this computer graphic graffiti led to the formulation of the VIDEOPLACE concept.

VIDEOPLACE is a computer graphic environment in which the participant sees his or her live image projected on a video screen. It may be alone on the screen, or there may be images of other people at different locations. In addition, there may be graphic objects and creatures which interact with the participant's image.

When people see their image displayed with a graphic object, they feel a universal and irresistible desire to reach out and touch it. (Fig. 2) Furthermore, they expect the act of touching to affect the graphic world. By placing each participant against a neutral background, it is possible to digitize the image of his silhouette and to recognize the moment when it touches a graphic object. The system can then cause the object to move, apparently in response to the participant's touch.

It is also possible for the computer to analyze the participant's image and to alter its appearance on the screen. By either analog or digital techniques, the participant's image can be scaled and rotated and placed anywhere on the screen. Thus, in principle, the participant could climb graphic mountains, swim in graphic seas, or defy gravity and float around the screen. The potential for new forms of interaction within this model is very rich, with certain application as an art form, likely application in education and telecommunication, as well as arguable application for general human-machine interaction.

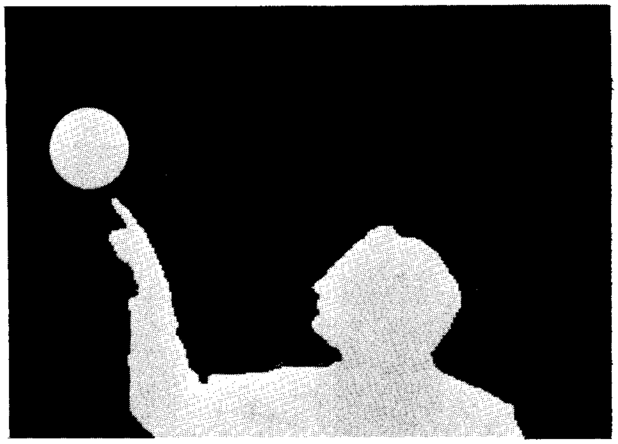

Fig. 2

Prototype System

A prototype VIDEOPLACE system has been constructed. Since understanding the movements of the participant's image appeared to be the most challenging issue, much of the initial effort was focussed on solving this problem. Graceful mechanisms for specifying and controlling the desired interactive relationships have been developed. To date, only very simple graphics have been used because of the very modest resources available and the fact that until recently commercial equipment did not emphasize high speed manipulation of raster data. To a great extent, we have worked with graphic hardware of our own construction which provides a number of features important to our interactions. In addition, we have recently acquired three Silicon Graphics workstations, which will greatly enhance our ability to create and manipulate realistic three-dimensional scenes.

The software employed to control the interactions is quite unusual. We believe the methodology is of general interest for graphics and other real-time applications. We treat the overall system as a model of a real-time intelligence. It is divided into two major components: the Cognitive System which runs on a VAX11/780 and the Reflex System which consists of a group of closely coupled dedicated processors operating on a specialized bus structure.

The Reflex System handles instantaneous decision making. The plan is for the Cognitive System to monitor the events in the Environment and the decisions of the Reflex System, in order to understand what is happening in semantic terms and then to make strategic decisions that will alter the future character of the interaction.

Although all of the communications are established, the Cognitive System is not yet performing this monitoring function. However, it has totally altered the programming process. Instead of writing a separate program for each interaction, we describe the desired causal relationships in conceptual terms. This conceptual representation is then translated into a form that the Reflex System can interpret in real-time. The long term objective is to develop an online real-time intelligence that understands the participant's behavior and the interaction in human terms.

CRITTER, A VIDEOPLACE Interaction

In one current interaction, the participant is joined by a single graphic creature on the screen. The behavior of this creature is very complex and context dependent. The intent is to produce the sensation of an intelligent and witty interaction between creature and the participant.

Initially, the creature sees the participant and chases his image about the screen. If the participant moves rapidly towards it, the creature, nicknamed CRITTER, moves to avoid contact. If the human holds out a hand, CRITTER will land on it and climb up the person's silhouette. As it climbs, its posture adapts to the contour of the human form. When it finally scales the person's head, it does a triumphant jig.

Once this immediate goal is reached, the creature considers the current orientation of the person's arms. If one of the hands is raised, it does a flying somersault and lands on that hand. If the hand is extended to the side but not above the horizontal, CRITTER dives off the head, rolls down the arm, grabs the finger and dangles from it. When the person shakes his hand, CRITTER falls off and dives to the bottom of the screen. Each time it climbs to the top of the participant's head, it is in a different state and is prepared to take a different set of actions. (Fig. 3a-h)

The CRITTER experience will soon be enhanced in a number of ways. Hardware has been built that shrinks the human image down to CRITTER size. The smaller size increases the number of relationships that can exist between the participant and the creature. Simple graphic scenes are being added. Both human and graphic entities will interact with these graphic props by moving among them, climbing them or hiding behind them. The new displays will provide a capability for three-dimensional scenery which can be navigated in real-time.

Practical Applications

The interface described is a deliberately informal one. The resemblance to video games might seem frivolous to the hard-nosed computer scientist used to catering to the needs of government agencies and three letter companies. However, games are a multi-billion dollar industry and by that measure practical. More importantly, games provide an extremely compelling interface whose advantages should be considered for more standard applications. Therefore, before adapting the techniques described to fit a more familiar practical context, we will examine their potential in the current VIDEOPLACE environment.

Computer Aided Instruction

In our culture, education is a sedentary activity imposed on naturally active creatures. Sti-

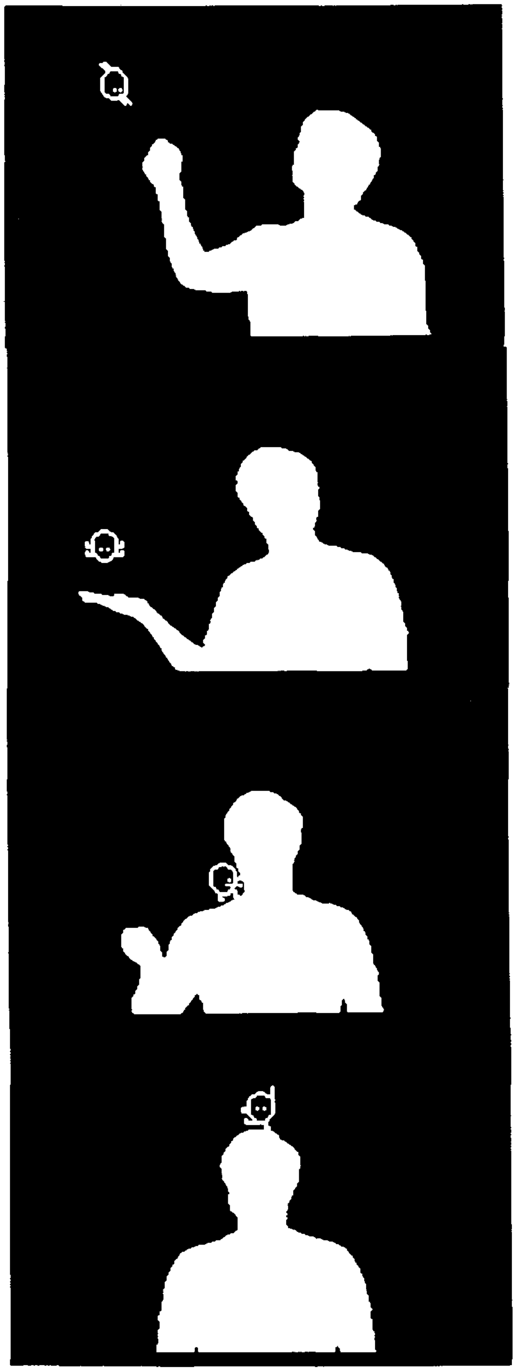

Fig. 3a-d

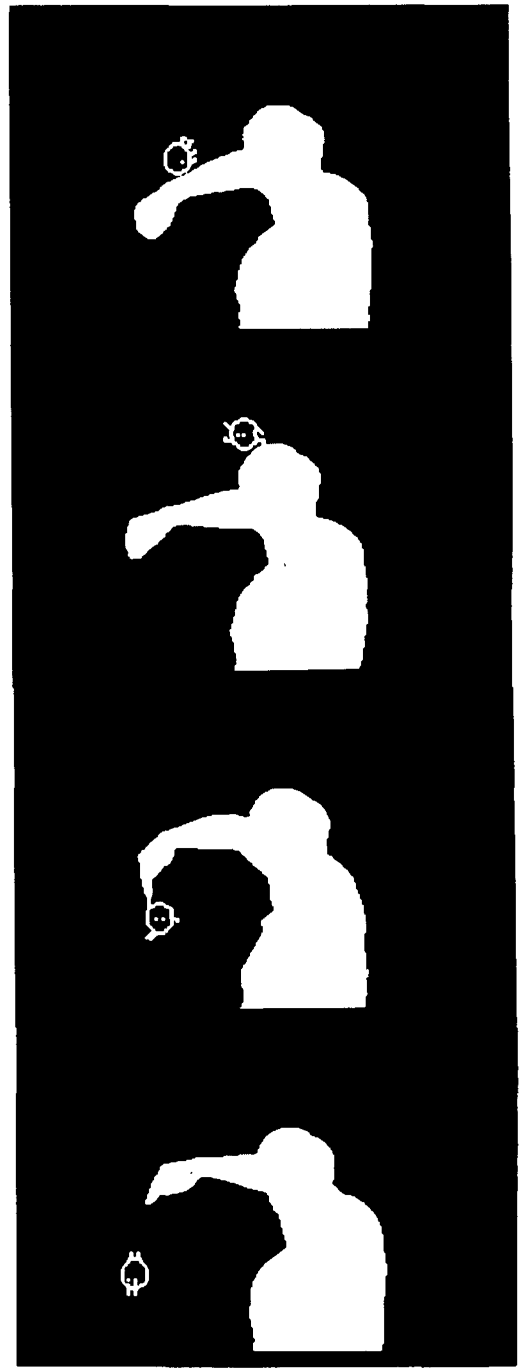

Fig. 3e-h

fling this energy is the first task of every elementary school teacher. As an alternative, VIDEOPLACE could be used to create a physically active form of Computer Aided Instruction in which the computer is used not to teach traditional material, but to alter what, as well as how, we teach.

In one proposal, which I first made formally to NSF in 1975, elementary school children were to be placed in the role of scientists landing on an alien planet. VIDEOPLACE would be used to define an artificial reality in which the laws of cause and effect are composed by the programmer. The task of the children would be to discover these laws. They would enter the Environment singly, interact with it and make individual observations of its rules. Under the guidance of their teacher, they would discuss their experiences and present their opinions. They would compare notes and formulate theories. Since each child would behave differently, in the VIDEOPLACE, individuals would have unique experiences and produce conflicting theories. They would argue and then revisit the Environment, executing critical experiments to resolve which theories were correct under which conditions. Thus, students would learn how observation leads to hypothesis formation, prediction, testing and reformulation. They would learn the process of scientific thought rather than memorizing vocabulary and performing mechanical calculations as they often do now.

Telecommunications

VIDEOPLACE was originally conceived and implemented as a telecommunications environment allowing people in different places to share a common video experience. While the possibility of such graphic interaction may seem unnecessary to communication, we should remember two points. First, a hundred years ago the telephone seemed to have no advantage over the telegraph which could transmit the content of messages equally well. Second, since communication between friends or business associates is not limited to words, it is clearly desirable to provide a place in which individuals who are geographically separated can share a common visual environment.

An example of this use of VIDEOPLACE is described in Artificial Reality (KRUE 83). A two-way computer graphic and live video telecommunications link was used to solve an engineering problem. In this experiment, the graphic images from two computers were viewed by television and combined by standard video techniques. Each participant pointed to the image on his local screen. The images of both of the participants' hands were combined with the graphic image, allowing them to gesture as naturally as if they were sitting together at a table. For the signal processing task at hand, the communication was complete.

Computing by Hand

A number of technologies are competing for space on the modern professional's desk. Telephones, answering devices, modems and computer terminals with touch screens are all candidates for the desk top. From the user's point of view, an empty desk is preferable. Two technology trends augur the removal of the computer terminal from the desk's surface. First, the keyboard will ultimately succumb to voice input. Second, flat screen displays of adequate resolution already exist. They are likely to be placed on a wall behind the desk, not on it, making touch screen input awkward.

The VIDEOPLACE techniques described in this paper can be used to duplicate any touch screen capability. A video camera pointed down at a desk surface can be used to create a VIDEODESK environment that will have several advantages over a touch screen.

In the VIDEOPLACE system, the user's hands can be used for any traditional graphics application. Since the system can detect when a person's hand touches a particular object, pointing and selection can be controlled. Similarly, a finger can be used to position the selected object in a design. A finger can also be used to draw on the screen, for example, to connect components in a logic design. We have already implemented simple menu selection, typing and finger painting systems. (Fig. 4)

Fig. 4

Video input offers more than a simple alternative to other pointing techniques. With the exception of the recent development of three-dimensional input devices, virtually all pointing devices are limited to two degrees of freedom. However, on the VIDEODESK, two hands can be used in concert to increase the user's bandwidth. In fact, in one common graphic application, it is easy to see the use for eight or more degrees of freedom. B-spline curves are used widely to design car bodies, ships hulls, turbine blades, etc. These curves are defined in terms of a relatively small number of control points. The user controls the shape of the curve by moving these points. With existing input devices only one point can be moved at a time. On the VIDEODESK, the tips of the index fingers and thumbs can be used to manipulate four control points simultaneously. (Fig. 5)

Conclusion

VIDEOPLACE is not so much a solution to existing problems, as an effort to stretch our thinking about the human-machine interface. We have already entered an era where most of the people using computers are no longer programmers in the traditional sense. We can look to a day when most of the people interacting with computers will not be users in the current sense.

Since computers are becoming less expensive than the people who use them, we can expect that as much computing power will be dedicated to providing a pleasing human-machine interface as is actually used to accomplish the user's application. As computer interaction becomes the dominant mode of performing work and transacting business, it becomes a significant ingredient in our quality of life. It is time to give the aesthetics of human-machine interaction serious thought.

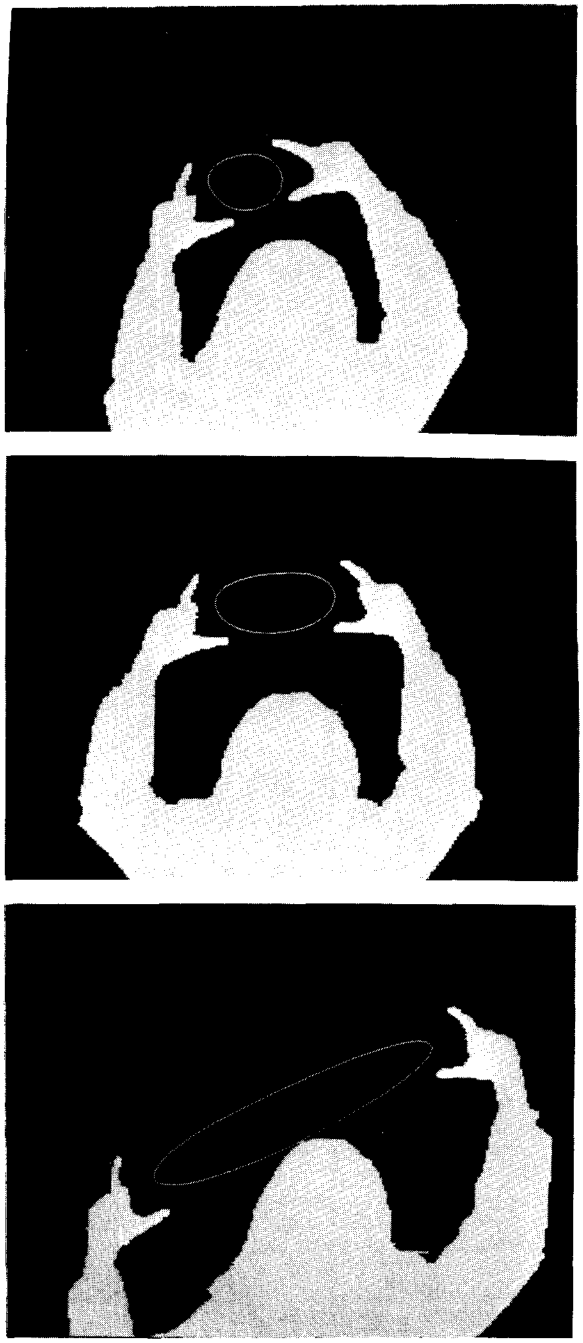

Fig. 5

Bibliography

[BATT72] Batter, James J. & Brooks, Frederick P. Jr., "GROPE-1," IFIPS 71, pp. 759-765.
[BOLT81] Bolt, Richard A., "Gaze Orchestrated Dynamic Windows." SIGGRAPH 81. pp. 32-42.
[BOLT80] Bolt, Richard A., "Put That There: Voice and Gesture at the Graphic Interfaces." SIGGRAPH 80. pp. 262-270.
[BOLT79] Bolt, Richard A., "Spatial Data Management." MIT 1979.
[CROC81] Crockett, D.W., "Triform Modules." CACM. Vol 24, No 6. pp. 344-350.
[CULL82] Cullingford, R.E., Krueger, M.W., Selfridge, M., "Automated Explanations in a CAD System." IEEE Transactions on SMC. January 1982.
[HAYE77] Hayes-Roth, F. and Lesser, V., "Focus of Attention in the Hearsay-II Speech Understanding System." IJCAI 77. pp. 27-35.
[KRUE83] Krueger, M.W., Artificial Reality. Addison-Wesley, 1983. 312 pp.
[KRUE81] Krueger, M.W., Cullingford, R.E. & Bellavance, D.A., "Control Issues in a CAD System with Expert Knowledge." Proceedings SMC. Oct. 1981.
[KRUE77] Krueger, M.W., "Responsive Environments." AFIPS 1977. 46:423-429.
[KRUE75] Krueger, M.W., IVAM, Annual Reports of NOAA Contract #5-315156 1975-1978.
[LIPP80] Lippman, Andrew. "Movie Maps: An Application of the Optical Disc to Computer Graphics." SIGGRAPH 81. pp. 109-120.
[SCHA77] Schank, R. and Abelson, R.P., "Scripts, Plans, Goals and Understanding." Erlbaum Press, 1977.
[SUTH68] Sutherland, Ivan. 1968. "A Head-Mounted Three-Dimensional Display." FJCC AFIPS 33-1:7575-764.