What Should You Wear to an Artificial Reality?

Myron W Krueger
Artificial Reality Corporation
Box 786, Vernon, CT 06066

Introduction

In the late 1960s, I was a graduate student in Computer Science at the University of Wisconsin. Unlike my peers, with engineering degrees, I had a liberal arts background. Thus, while they were interested in the efficiency of the computer, I thought of its philosophical significance.

Specifically, I saw the encounter between human and machine as the central drama of our time. Therefore, I was interested in working on the human-machine interface. The relationship between human and machine is a permanent part of the human condition, as fundamental as particle physics. Thus, the human-computer interface is a fundamental research problem that should be considered from the broadest possible perspective. Unfortunately, while there was a great deal of posturing in the computer science community about being scientific, at that time, the development of computers was completely device centered. Researchers focussed on the design of the computer and its operating software. The human interface was onlya thin veneer applied to the computer to make it a little easier to use.

It seemed apparent to me that the human interface consisted of two components: the computer and the human. The former was changing more rapidly than any technology in history while the latter was not changing at all. Therefore, the research community should have focussed its efforts on immutable human nature rather than attempting to track every twist and turn of technology development. I felt that the ultimate computer interface was the human body and the human voice. Instead of receiving input from us through a 100 year old keyboard, the computer should perceive our movements and understand our speech.

Somewhat earlier, Ivan Sutherland had suggested that the ultimate computer interface would allow you to sit within a computer graphic room and look around it. I did not know this at the time, but if I had, I would have objected strongly. What had always bothered me about computers specifically, and intellectual

activity in general, was that you had to sit down to do it. I wanted my body back. Thus, I always thought of creating an interface that would permit me to move around as I interacted with the computer.

Although my original goals were technical, I quicky realized that the scientific method was not appropriate for discovering the ideal design for this interface. I reasoned that if we were going to use the computer all day, every day, we would have an unpleasant life if we had an unpleasant interface. Thus, I saw that the design of such an interface had an important aesthetic component. I felt that the computer interface should be judged as much by how it pleased the user as by its efficiency.

In considering what the ultimate interface should be, I decided to study the people who had the best relationship to their tools: artists and musicians. Interacting with these people exposed me to the Experiments in Art and Technology movement and the early efforts in computer art. Up until that time, I had had very little interest in art, but suddenly I knew why. What was being presented as art in the twentieth century was universally accomplished with antique tools. As I examined the activities of artists, I observed that many used technology—just obsolete technology. I realized that my lack of interest was based on the fact that traditional media had frozen art in the past. Historically, artists were versed in technology and were the agents of technological change if their work demanded it. Certainly, lost wax casting in ancient Greece was high tech. The assumption that an artist must be a technological incompetent is a very recent invention.

At the same time, when I examined contemporary art and technology or computer art, I disagreed with the basic premise of the work. The goal seemed to be to use the new technology to create traditional works of art. It seemed to me that if there was such a thing as computer art, that it should represent at least as great a change in art as the computer had engendered in other fields—for example, accounting. When I thought about what was unique about the opportunity that the computer offered, I decided that the fundamentally new ingredient was interactivity. A computer-controlled art work could perceive its viewer and respond to his behaviour in real time. Response was the medium! I was absolutely convinced that this conclusion was correct, which surprised me since I had not concerned myself about the nature of art previously. I discovered that I was very excited about the possibility of such art and would care very much about it if it were created. To my surprise, I found that traditional artists

had absolutely no interest in my vision. If I was ever to see the art that I wanted, I would have to create it. I would have to become an artist.

In 1969, I did a collaborative work with a group of artists and technologists. When I found that they resisted making interactivity the focus of the work, I vowed to create work in which composed interaction was the primary goal.

In 1970, I created an installation called METAPLAY which was an "experiment" in the old-fashioned, romantic sense of the word. Today, scientists always have to know the result that their experiments will yield before they apply for funding let alone run the experiment. In METAPLAY, I did an experiment to see what would happen. I wanted to discover the innate expectations of the technologically naive human. How could a person be induced to interact with a computer without verbal instruction?

I knew that I wanted to respond to a person's behavior using the computer. This desire implied that the computer would be able to perceive the participant's behavior. I had no idea how to achieve that in 1970. Therefore, I decided to fake it. Like the Wizard of Oz, I created my desired interface with smoke and mirrors—a bit of interactive theater. I used the only mechanism that I was aware of that could understand human behavior in real time—my own eyes.

People walked into the Union Gallery at the University of Wisconsin. They were confronted with a 2.7 meter by 3.3 meter rear-projection video screen that contained their own life-sized image. They saw grafitti being drawn on their image. I was sitting a mile away in the computer center drawing with a data tablet. My drawing appeared on an Adage vector graphic display. A video camera was aimed at the Adage screen and its signal was transmitted across campus to the gallery where it was superimposed on the live image and the resulting composite image was projected in front of the participants. In addition, the composite image was transmitted back to the computer center where it was displayed on a monitor that provided the feedback I needed draw.

People were mildly entertained by the grafitti for a while. Then, they started to make the situation interactive. When they saw the cursor coming toward them, they would run away from it. They would bat it away as if it were alive. Based on that reaction, I changed the cursor from a cross to a circle which I placed at the top of the screen when people entered. Inevitably, people wondered what would happen if they touched the image of the ball with the image of their hand. They would reach up and hit the ball. I would move it across the screen as if in

response to their blow. Typically, a person on the other side of the screen would reach up to hit it back. I had people playing vollyball with a nonexistent ball in 1970.

Then, one day I was drawing something on a person's hand. He was embarrassed by the attention and moved his hand. I followed his hand with the pen. The result was that a line appeared on the screen. It looked as if he had drawn the line with his finger. He moved his finger again and I followed. After a moment of crudely successful drawing, the person next to him reached up and took the line away with his finger. We were able to lead thousands of participants to discover the drawing interaction without saying a word to them. This process was repeated ten hours a day, seven days a week, for six weeks. Many interactions were invented, revealing a set of expectations that were established by juxtaposing a person's image with a graphic object.

One day we were working on a technical problem. We were sending data from a sensory floor in the gallery to the graphics computer in the computer center. There was a problem in data transmission. We talked on the phone trying to identify the nature of the problem and then its source. It occurred to me that we had a better way of discussing our problem. I asked my colleague to display the information he was transmitting on the computer in the gallery, which had a limited graphic capability. Next, I had him aim the video camera at that computer's screen. The result was that the composite video image showed the data transmitted from the gallery computer at the top half of the screen and the data received by the computer center computer at the bottom half of the screen. As we discussed the composite image, it was natural for us to move our hands in front of our respective cameras to point at features we were discussing. It was just as if we were sitting together on the same side of a table talking about information on a piece of paper. The illusion of being together was compelling enough that when the image of my hand briefly overlapped the image of his, he moved his hand. I barely noticed the first time this happened. After several occurrences, I realized that he was unwilling to hold hands with me. Personal distance was being observed in this purely visual world.

Subsequent experimentation confirmed my suspicion. For most people, their image is an extension of themselves. What happens to their image, happens to them. What touches their image, they feel. Thus, I realized that I had stumbled onto a very important set of principles. While we usually think of telecommunication as being between two points, in METAPLAY, we had created a

new place which consisted of the information that the two remote participants shared simultaneously. In the playful part of METAPLAY, one party communicated through computer graphics and the other communicated through physical gesture. In the serendipitous teleconference, we both used our hands and voices to interact. (I should acknowledge that the only reason that I had resorted to telecommunication in the first place was that the university would not let me move the graphics computer to the gallery.)

The ultimate extension of these concepts was the creation of a new telecommunication place that I dubbed VIDEOPLACE. One would enter VIDEOPLACE through an installation similar to that used in METAPLAY. Instead of being alone on the screen, your image would be joined by the images of other people who were in distant locations. The experience would be similar to METAPLAY, except that what had been accomplished with human intervention would instead be accomplished automatically by computer. To make the remote participants look like they were in a new place, the video image would include graphic scenery which would provide a context for their interaction. Since we had already discovered that people who see themselves together, act like they are together, and that they expect something to happen when their images touch graphic objects, we decided to make that sense of touch effective. In VIDEOPLACE, when a person touches the image of another person or the image of a graphic object, the computer can detect that contact and respond to it in any imaginable way, including moving the object, making it disappear, changing its color, or causing it to emit a tone.

In addition, the participant's image can be subjected to forces that are not operating on his body. The computer can move, scale, and rotate the participant's live video image exactly as if it were a graphic object. This movement can be in response to the participant's action, in response to the action of another participant or graphic creature, or at a scripted moment in the interaction. The result is that VIDEOPLACE is an artificial reality in which the laws of cause and effect are composed from moment to moment.

VIDEOPLACE was proposed as the theme of the US Bicentenniel. It was to be created on a worldwide scale. There were to be VIDEOPLACE installations at several locations in the US, Europe, and Japan. While VIDEOPLACE was obviously not implemented on such a global scale, it was an official Bicentenniel project. It did receive some funding from the National Endowement for the Arts, and a preliminary version of VIDEOPLACE was exhibited in the Milwaukee Art Museum in 1975.

In 1974, I began the construction of the VIDEOPLACE system. While no components from that era still exist, the system has been in constant operation since that period. It slowly metamorphoses from one technology to the next. The organic development of the system is unlike any technology project I am aware of. By the mid-1970's, VIDEOPLACE was capable of rudimentary operation. In the following decade, the system grew by accretion with the gradual improvement of its ability to understand what the participant standing in front of the camera was doing with respect to the graphic world. In the early 1980s, VIDEOPLACE had evolved to the point where it was capable of a wide range of interactions based on real-time understanding of the participant's behavior.

At the end of 1982, my 1974 dissertation appeared in book form as Artificial Reality (Addison-Wesley). It considered the technology and aesthetics of interactive media like VIDEOPLACE. It also forecasted the current goggle-and-glove technology as another way of accomplishing similar results.

In 1985, I went public with a paper describing working technology at the SIGCHI Conference (Computer-Human Interaction). Later that year, VIDEOPLACE was shown in the SIGGRAPH Art Show, also in San Francisco. The VIDEOPLACE exhibited at that time was a single-ended installation in which one or more people interacted with the computer. They stood before a neutral background of translucent plastic, backlit by fluorescent lights to provide a very high contrast image which enables the computer to distinguish the participant from the background. This video image was thresholded and subjected to a two-dimensional filter to eliminate noise pixels. The resulting binary image is passed on to a series of specialized computers that operate in parallel to analyze the participant's sihouette in order to locate head, hands, torso, and fingers. These processors are thousands of times faster than a PC. Since they operate in parallel with the host computer, the elaborate perception algorithms exact no penalty from the performance of the system.

The result is not a single piece in the traditional artistic sense, but a new medium in which many interactive vignettes have been composed. There are many reasons that I have chosen to develop the art form in this fashion. Many of these are direct attacks on the comfortable tradition of art which conforms to old norms but offers nothing to the world we live in.

I have long observed that people consume art in what I term "supermarket mode."

They visit galleries and museums expecting to expose themselves to hundreds of works of art per hour. Obviously, they cannot truly see this much art or, if they can, it is an incredibly weakened form of art that is being experienced. To thwart this process, I always show many types of work, done for a variety of motivations, in a variety of styles. The effect of this is that when one of these art "shoppers" passes by, they look in and see a particular type of activity going on and rush on, desperate to fill their quota. Later, if they chance to compare what they saw with the experience of another person, they discover that they each saw completely different types of work. The result is that they feel compelled to return to really see the work. This is my "second strike" capability. It has been a conscious strategy since METAPLAY and there are always a few individuals who confide that they have had exactly this experience.

Another reason I have pursued this strategy has been that I have always been struck by the number of modern artists whose style was so consistent that their work could be recognized at one hundred yards. Their entire careers can be distilled down to a single artistic statement—an aesthetic one liner. Rather than be trapped into a single style of work that would have made life easy for critics, I have deliberately pursued a wide variety of approaches to celebrate the rich opportunity that interactive realities offer.

Artificial realities are different than the traditional museum arts of painting and sculpture in that they are composed through time and involve motion. Thus, they can be more like film and theater than the traditional static forms. These differences from existing forms disturb traditionalists who believe that the limitations of the traditional arts are virtues. In response, I offer the following thought experiment. If the cave painters of Lascaux had been offered animated paint, would they have used it? If they could have made their bison behave in realistic ways, would they have done so?

My own interest in artificial realities lies in the possibility that they will yield totally new kinds of aesthetic experience. If they produce nothing more provocative than contemporary art then they are not worth doing. Therefore, I pursue the directions that seem most fundamental to the medium. Often, these are exactly the directions that are most unlike traditional art.

As an interactive artistic medium, artificial realities differ from earlier art in that the artist leaves the work unfinished. Its completion is provided by the unique behavior of each individual who interacts with it. This possibility confuses the identity of the artist. In the past, humanity could afford very little

creative effort. Therefore, a few individuals were assigned the role of aesthetic specialists to do art for the rest of humanity who were too busy plowing the fields. Today, however, there is much more leisure and the traditional role of the artist reveals an unseemly side. Since artists alone had the time to develop aesthetic skills, their work was admired by people who knew they could not duplicate it. Thus, their work intimidated the layman and stiffled his creative instincts. He could not paint because he was not an artist. He could not compose because he was not a musician. However, today, it is appropriate to invite all individuals to quicken their aesthetic sensibilities, not to consume art as passive spectators but to actively create it.

One class of VIDEOPLACE interactions pursues this goal. The participant enters the environment and finds that his movements create aesthetically pleasing patterns. He has been given a new art medium to play with. Since he does not recognize this medium as corresponding to any existing art form, he is not inhibited about exploring it. Therefore, he will often move his body motivated solely by aesthetic goals, even though aesthetics ordinarily would never guide his behavior. What is important is that the aesthetic experience is created by the person who will appreciate it. It is not created by one specialist to be critiqued by another specialist who will then instruct the public as to what they should like and why.

The most basic of this class of interactions was based on the METAPLAY drawing interaction. In DIGITAL DRAWING, you hold up your finger and the image of your finger draws on the screen. It is liberating to be able to draw without touching anything. It is especially charming if you try to erase what you have drawn by rubbing your hand over it. The computer recognizes the gesture and erases the screen as if by magic. In a variation on this interaction, you can draw an invisible path that is continually traversed by flowing particles. By synchronizing your drawing movements with the flow of the particles, you can create a complex choreography of light. In related interactions, you can paint with your body or just your fingertips. You can create abstract geometric shapes or become an animated Mondrian painting. In a more sophisticated interaction, eight copies of your shrunken image are arranged on the screen in a circular pattern. The dynamics of the resulting kaleidoscopic medium are so rich that it will often captivate an individual for an hour.

The CRITTER interaction was the result of a tremendous effort. CRITTER is a simple computer-generated creature who perceives the participant's behavior

and interacts with him in a playful way. If the participant moves, CRITTER will chase him. If the participant holds out his hand, CRITTER will land on it. If the participant holds still, CRITTER will climb to the top of his head and do a little jig in celebration upon reaching the summit. It is also possible to get CRITTER to dangle from your finger. If you shake your finger, CRITTER will fall off and plunge to the bottom of the screen where it flattens as it strikes. After observing thousands of people attempting to capture CRITTER, we have made it possible. The first step is to surround it with your image. CRITTER will try to escape frantically. If the surrounding area is slowly constricted, it will explode—only to be resurrected a moment later. CRITTER is a stand-in for artificial intelligence or, at least, artificial personality. It is a playful artificial entity that is so engaging that you forget that the creature you are playing with is not really there.

Another theme of the work has resulted from the manner in which it was originally developed. I thought of METAPLAY as a way of simulating a computer-human interface that I could not yet implement. Success would occur when I could replace myself by a computer. What really happened was that I had stumbled onto a superior concept—computer meditated human-to-human interaction. The computer was allowing me to interact with another person in a new way. I could establish a simple world in which two people in different locations could interact with each other. This was the first time that telecommunication was identified as a fundamental part of the artificial reality concept. The artistic opportunity was that the artist could define the circumstance in which two people would interact and could thus be a third party to their interaction.

Towards this goal, I connected the VIDEOPLACE environment with the VIDEODESK, which has an overhead camera aimed down at a seated person's hands as they rest upon a light table. The image of the VIDEODESK participant's hands can then appear in a composite image that includes the VIDEOPLACE participant's full body. Either image can be moved, scaled, or rotated as a consequence of actions in the graphic world. The VIDEODESK participant controls the selection of response rules with her left hand. Since she has more knowledge and power that the VIDEOPLACE participant, her role is to facilitate a playful experience for the other participant. She is a benign Big Mother, as opposed to her tyrannical sibling.

She can join the other participant in any of the single-ended interactions,

making them into cooperative aesthetic teleconferences. She can also take advantage of the juxtaposition of her image with the other participant's. There is a tension created when the two images touch. In one case, we take advantage of this by making a sound when the VIDEODESK participant's finger touches the image of the VIDEOPLACE participant. A different sound is made when the VIDEOPLACE participant touches her hands with his finger. Thus, while there is an embarrassment about touching, nothing happens unless one participant touches the other.

In another variation, the VIDEODESK hand reaches in and pushes the VIDEOPLACE participant's image across the screen. Seeing their image abused in this way usually leads the latter to push back. Often a full scale fight ensues. Then, the VIDEODESK hand strikes the VIDEOPLACE participant's image in the head. The image is knocked down and pops back up again—like the inflatable children's toy with sand in the bottom.

In another case, the VIDEODESK hand extends a finger and a graphic string dangles from its tip. At the end of the string hangs the shrunken image of the VIDEOPLACE participant. Finding himself suspended in this way, this participant wonders if there is any way he can make his image swing on the string. He tries a few exploratory moves. He discovers that by moving from side to side he can make his image swing, and then by timing his movements correctly, he can swing higher and higher. He wonders, can he do a 360? After considerable, frantic effort, he succeeds.

The interactions I just described are between the VIDEOPLACE and the VIDEODESK participants. However, there will often be a small crowd of people watching. They will cheer when the VIDEOPLACE participant succeeds in making his image swing all the way around. They will also laugh when his image is knocked down. How often have you ever seen a computer make people laugh?

What relationship does the VIDEOPLACE medium have to the goggle-and-glove technology that is often considered the orthodox way of creating artificial realities? First, the instrumented glove is a way for the computer to see what the participant's hand is doing. The instrumented suit is a way of perceiving the articulation of the whole body. Video cameras are a far less obtrusive way of accomplishing the same results. While the VIDEOPLACE system currently only understands the movements of the participant's two-dimensional silhouette, using additional cameras to understand the participant's movements in three dimensions is a straightforward, if difficult, problem. As long as the glove or

the suit do not provide tactile feedback, there is no great advantage to them.

Similarly, the purpose of goggles is to completely immerse the participant in a computer-generated world. Admittedly, one video projector does not do that. However, it would be quite easy to use several projectors to surround the participant. Alternatively, light-weight glasses, recently designed for viewing 3D movies, could be worn. These glasses block one eye and then the other as the appropriate stereoptic view of the scene is projected. These glasses are far less encumbering than current goggles.

In the long run, the reality goggles will become less encumbering. Eventually, they will be the same size as standard eyeglasses or even contact lenses. With such a convenient display, wearing instrumented clothing will be even less attractive. Our expectation is that when such goggles are available, it will be natural to use video cameras to perceive the participant's body. Thus, the VIDEOPLACE medium may be combined with the head-mounted display. One advantage of this approach is that the person's facial expression will be visible to the video camera and thus their image in the three dimensional graphic world can also have a face. With the current goggle-and-glove paradigm, there is simply no way for the person's facial expressions to be captured or displayed.

Currently, the VIDEOPLACE system responds far more rapidly than the goggle-and-glove systems, in spite of the fact that analyzing a person's image through a video camera in 1/30 second is a more difficult problem than that posed by monitoring bend sensors worn on the participant's joints. The reason is that the VIDEOPLACE system has 12 speciallized processors dedicated to perceiving human images that operate 1000 times faster than a personal computer.

The difference in performance is based on an important philosophical difference between the implementers of the different systems. An interactive experience is compelling because it responds in interesting ways to the participant's behavior. Its authority does not derive from the realism of the graphic world portrayed. Similarly, theater succeeds because of the dramatic human interactions depicted, not primarily because of the stage craft. Indeed, there are types of theater in which sets are not employed. Of course, graphic realism will be of interest in the future. However, it is currently achieved by sacrificing the immediacy of the experience. Thus, when you are wearing goggles and turn your head, there is a very noticeable lag (sometimes as long as a second) before the graphic world catches up with you. When you move your hand, your graphic hand

hangs in the air for a long moment before it moves. These lags are very disturbing. They instantly alert you to the fact that the experience is not real. They are also avoidable. Simple graphic worlds can be rendered more rapidly than complex ones.

By way of contrast, the VIDEOPLACE participant is represented by his silhouette image, not his full-color video image. However, the participant never asks, "Is that me?" Instead, when they see CRITTER approach their silhouette's head, they duck. In a different example, consider the STEP LIGHTLY piece that I did for the SIGGRAPH Art Show in 1990. A laser mounted on the ceiling drew simple graphics on the floor that responded to participants' movements around the room. The most successful interaction showed a fish that swam after participants as they walked about. The graphic fish was very crude, but no one complained that it did not look like a fish. Its reaction to their movements was so immediate that its behaviour was realistic, even if its appearance was not. Visitors reacted instantaneously to the reality of the experience by running away and laughing—the experience was very funny.

The primary concern should be what happens in the artificial reality, not how it looks. Whether used for practical, playful, or aesthetic purposes, there are technical constraints that define the medium. If these are not observed, you are not simulating an artificial reality. Movies are always shown at 24 frames per second. The frame rate is not up to the discretion of the filmmaker. Interactive experiences require at least as high a frame rate.

Is the fact that artificial reality is an art form important? What about practical applications? Are they not more important? We have a blindness about the issue of practicality. Consider television. If television had been invented by NASA, they would have had to emphasize that it could be used for applications in space, in medicine, and in education. These claims would be accurate. However, these practical applications pale beside the culture-defining significance of the entertainment medium that television has proved itself to be. Artificial realities are a new realm of human experience, and as such, their invention is as significant as that of theater, the novel, film, or television.