

# VIRTUAL ENVIRONMENTS, PERSONAL SIMULATION & TELEPRESENCE

Scott S. Fisher

## 1. Media Technology and Simulation of First-Person experience

"Watch out for a remarkable new process called SENSORAMA! It attempts to engulf the viewer in the stimuli of reality. Viewing of the color stereo film is replete with binaural sound, colors, winds, and vibration. The original scene is recreated with remarkable fidelity. At this time, the system comes closer to duplicating reality than any other system we have seen!"<sup>1</sup>

For most people, "duplicating reality" is an assumed, if not obvious goal for any contemporary imaging technology. The proof of the "ideal" picture is not being able to discern object from representation — to be convinced that one is looking at the real thing. At best, this Judgement is usually based on a first order evaluation of "ease of identification"; i. e. realistic pictures should resemble what they represent. But resemblance is only part of the effect. In summing up prevailing theories on realism in images, Perkins comments: "Pictures inform by packaging information in light in essentially the same form that real objects and scenes package it and the perceiver unwraps that package in essentially the same way."<sup>2</sup>

What is most limited in contemporary media is the literal process involved in "unwrapping" the image. Evaluation of image realism should also be based on how closely the presentation medium can simulate dynamic, multi-modal perception in the real world. A truly informative picture, in addition to merely being an informational surrogate, would duplicate the physicality of confronting the real scene that it is meant to represent. The image would move beyond simple photo-realism to immerse the viewer in an interactive, multi-sensory display environment. Methods to implement and evaluate these interdependent factors contributing to image realism lie in the emerging domain of Media Technology. Until recently, significant developments in this area have usually been dictated by economics, available technology and, as mentioned, cursory ideas about what types of information are sufficient in image representation. For example, the medium of television, as most experience it, plays to a passive audience. It has little to do with the nominal ability to "see at a distance" other than in a vicarious sense; it offers only interpretations of remote events as seen through the eyes of others with no capability for viewpoint control or personal exploration. And, although this second hand information may be better than no information at all, a "first-person", interactive point of view can offer added dimensions of experience:

"We obtain raw, direct information in the process of interacting with the situations we encounter. Rarely intensive, direct experience has the advantage of coming through the totality of our internal processes—conscious, unconscious, visceral and mental — and is most completely tested and evaluated by our nature. Processed, digested, abstracted second-hand knowledge is often more generalized and concentrated, but usually affects us only intellectually — lacking the balance and completeness of experienced situations ... Although we are existing more and more in the realms of abstract, generalized concepts and principles, our roots are in direct experience on many levels, as is most of our ability to consciously and unconsciously evaluate information."<sup>3</sup>

In the past few decades, changing trends in Media Technology have begun to yield innovative ways to represent first-person or "direct experience" through the development of multi-sensory media environments in which the viewer can interact with the information presented as they would in encountering the original scene. A key feature of these display systems (and of more expensive simulation systems) is that the viewer's movements are non-programmed; that is, they are free to choose their own path through available information rather than remain

restricted to passively watching a "guided-tour". For these systems to operate effectively, a comprehensive information database must be available to allow the user sufficient points of view. The main objective is to liberate the user to move around in a virtual environment, or, on a smaller scale, to viscerally peruse a scene that may be remotely sensed or synthetically generated. In essence, the viewer's access to greater than one viewpoint of a given scene allows them to synthesize a strong visual percept from many points of view; the availability of multiple points of view places an object in context and thereby animates its meaning.

## **2. The Evolution of Virtual Environments**

Matching visual display technology as closely as possible to human cognitive and sensory capabilities in order to better represent "direct experience" has been a major objective in the arts, research, and industry for decades. A familiar example is the development of stereoscopic movies in the early 50's, in which a perception of depth was created by presenting a slightly different image to each eye of the viewer. In competition with stereo during the same era was Cinerama, which involved three different projectors presenting a wide field of view display to the audience; by extending the size of the projected image, the viewer's peripheral field of view was also engaged. More recently, the omnimax projection system further expands the panoramic experience by situating the audience under a huge hemispherical dome onto which a high-resolution, predistorted film image is projected; the audience is now almost immersed in a gigantic image surround.

In 1962, the "Sensorama" display previously noted was a remarkable attempt at simulating personal experience of several real environments using state of the art media technology. The system was an elegant prototype of an arcade game designed by Morton Heilig: one of the first examples of a multi-sensory simulation environment that provided more than just visual input. When you put your head up to a binocular viewing optics system, you saw a first-person viewpoint, stereo film loop of a motorcycle ride through New York City and you heard three-dimensional binaural sound that gave you sounds of the city of New York and of the motorcycle moving through it. As you leaned your arms on the handlebar platform built into the prototype and sat in the seat, simulated vibration cues were presented. The prototype also had a fan for wind simulation that combined with a chemical smellbank to blow simulated smells in the viewer's face. As an environmental simulation, the Sensorama display was one of the first steps toward duplicating a viewer's act of confronting a real scene. The user is totally immersed in an information booth designed to imitate the mode of exploration while the scene is imaged simultaneously through several senses.

The idea of sitting inside an image has been used in the field of aerospace simulation for many decades to train pilots and astronauts to safely control complex, expensive vehicles through simulated mission environments. Recently, this technology has been adapted for entertainment and educational use. "Tour of the Universe" in Toronto and "Star Tours" at Disneyland are among the first entertainment applications of simulation technology and virtual display environments; About 40 people sit in a room on top of a motion platform that moves in synch with a computer-generated and model-based image display of a ride through a simulated universe.

This technology has been moving gradually toward lower cost "personal simulation" environments in which the viewer is also able to control their own viewpoint or motion through a virtual environment — an important capability missing from the Sensorama prototype. An early example of this is the Aspen Movie Map, done by the M.I.T. Architecture Machine Group in the late 70's.<sup>4</sup> Imagery of the town of Aspen, Colorado was shot with a special camera system mounted on top of a car, filming down every street and around every corner in town, combined with shots above town from cranes, helicopters and airplanes and also with shots inside buildings. The Movie Map gave the operators the capability of sitting in

front of a touch-sensitive display screen and driving through the town of Aspen at their own rate, taking any route they chose, by touching the screen, indicating what turns they wanted to make, and what buildings they wanted to enter. In one configuration, this was set up so that the operator was surrounded by front, back, and side-looking camera imagery so that they were completely immersed in a virtual representation of the town.

Conceptual versions of the ultimate sensory-matched virtual environment have been described by science fiction writers for many decades. One concept has been called "telepresence", a technology that would allow remotely situated operators to receive enough sensory feedback to feel like they are really at a remote location and are able to do different kinds of tasks. Arthur Clarke has described "personalized television safaris" in which the operator could virtually explore remote environments without danger or discomfort. Heinlein's "waldos" were similar, but were able to exaggerate certain sensory capabilities so that the operator could, for example, control a huge robot. Since 1950, technology has gradually been developed to make telepresence a reality.

Historically, one of the first attempts at developing these telepresence visual systems was done by the Philco Corporation in 1958. With this system an operator could see an image from a remote camera on a CRT mounted on his head in front of his eyes and could control the camera's viewpoint by moving his head.<sup>5</sup> A variation of the head-mounted display concept was done by Ivan Sutherland at MIT in the late 60's.<sup>6</sup> This helmet-mounted display had a see-through capability so that computer-generated graphics could be viewed superimposed onto the real environment. As the viewer moved around, those objects would appear to be stable within that real environment and could be manipulated with various input devices that they also developed. Research continues at other laboratories such as NASA Ames in California, the Naval Ocean Systems Center in Hawaii and MITI's Tele-existence Project in Japan: Here the driving application is the need to develop improved systems for humans to operate safely and effectively in hazardous environments such as undersea or outer space.

### **3. VIEW: The NASA / Ames Virtual Environment Workstation**

In the Aerospace Human Factors Research Division of NASA's Ames Research Center, an interactive Virtual Interface Environment Workstation (VIEW) has been developed as a new kind of media-based display and control environment that is closely matched to human sensory and cognitive capabilities. The VIEW system provides a virtual auditory and stereoscopic image surround that is responsive to inputs from the operator's position, voice and gestures. As a low cost, multipurpose simulation device, this variable interface configuration allows an operator to virtually explore a 360 degree synthesized or remotely sensed environment and viscerally interact with its components.<sup>7-11</sup>

The current Virtual Interface Environment Workstation system consists of: a wide-angle stereoscopic display unit, glove-like devices for multiple degree-of-freedom tactile input, connected speech recognition technology, gesture tracking devices, 3D auditory display and speech-synthesis technology, and computer graphic and video image generation equipment.



Figure 1

When combined with magnetic head and limb position tracking technology, the head-coupled display presents visual and auditory imagery that appears to completely surround the user in 3-space. The gloves provide interactive manipulation of virtual objects in virtual environments that are either synthesized with 3D computer-generated imagery, or that are remotely sensed by user-controlled, stereoscopic video camera configurations. The computer image system enables high performance, realtime 3D graphics presentation that is generated at rates up to 30 frames per second as required to update image viewpoints in coordination with head and limb motion. Dual independent, synchronized display channels are implemented to present disparate imagery to each eye of the viewer for true stereoscopic depth cues. For realtime video input of remote environments, two miniature CCD video cameras are used to provide stereoscopic imagery. Development and evaluation of several head-coupled, remote camera platform and gimbal prototypes is in progress to determine optimal hardware and control configurations for remotely controlled camera systems. Research efforts also include the development of realtime signal processing technology to combine multiple video sources with computer generated imagery.

#### 4. Virtual Environment Applications

Application areas of the virtual interface environment research at NASA Ames are focused in two main areas — Telepresence and Dataspace:

TELEPRESENCE-The VIEW system is currently used to interact with a simulated telerobotic task environment. The system operator can call up multiple images of the remote task environment that represent viewpoints from free-flying or telerobot-mounted camera platforms. Three-dimensional sound cues give distance and direction information for proximate objects and events. Switching to telepresence control mode, the operator's wide-angle, stereoscopic display is directly linked to the telerobot 3D camera system for precise viewpoint control. Using the tactile input glove technology and speech commands, the operator directly controls the robot arm and dexterous end effector which appear to be spatially correspondent with his own arm.

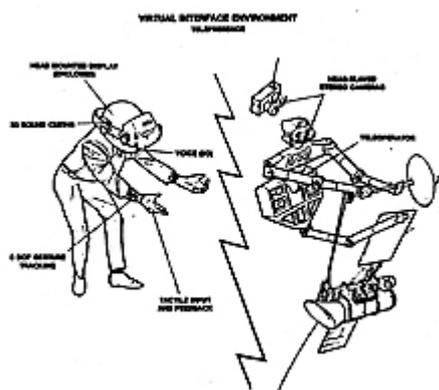


Figure 2

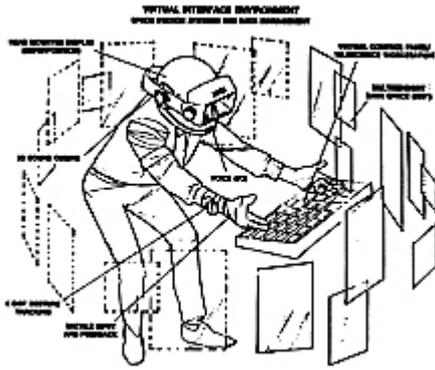


Figure 3

DATASPACE-Advanced data display and manipulation concepts for information management are being developed with the VIEW system technology. Current efforts include use of the system to create a display environment in which data manipulation and system monitoring tasks are organized in virtual display space around the operator. Through speech and gesture interaction with the virtual display, the operator can rapidly call up or delete information windows and reposition them in 3-space. Three-dimensional sound cues and speechsynthesis technologies are used to enhance the operators overall situational awareness of the virtual data environment. The system also has the capability to display reconfigurable, virtual control panels that respond to glove-like tactile input devices worn by the operator.

### 5. Personal Simulation: Architecture, Medicine, Entertainment

In addition to remote manipulation and information management tasks the VIEW system also may be a viable interface for several commercial applications. So far, the system has been used to develop simple architectural simulations that enable the operator to design a very small 3D model of a space, and then, using a glove gesture, scale the model to life size allowing the architect / operator to literally walk around in the designed space. Seismic data, molecular models, and meteorological data are other examples of multidimensional data that may be better understood through representation and interaction in a Virtual Environment.

Another Virtual Environment scenario in progress involves the development of a surgical simulator for medical students and plastic surgeons that could be used much as a flight simulator is used to train jet pilots. Where the pilot can literally explore situations that would be dangerous to encounter in the real world, surgeons can use a simulated "electronic cadaver" to do pre-operation planning and patient analysis. The system is also set up in such a way that surgical students can look through the eyes of a senior surgeon and see a first-person view of the way he or she is doing a particular procedure. As illustrated in the following figure, the surgeon can be surrounded with the kinds of information windows that are typically seen in an operating room in the form of monitors displaying life support status information and X-rays.

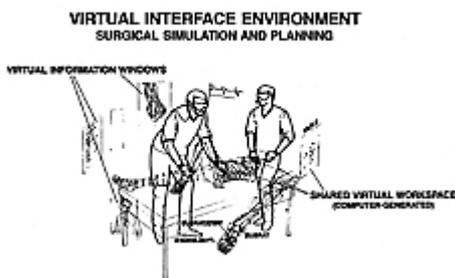


Figure 4

Entertainment and educational applications of this technology could be developed through this ability to simulate a wide range of real or fantasy environments with almost infinite possibilities of scale and extent. The user can be immersed in a 360-degree fantasy adventure game as easily as he or she can viscerally explore a virtual 3D model of the solar system or use a three-dimensional paint system to create virtual environments for others to explore.

## 6. Tele-Collaboration through Virtual Presence

A major near-term goal for the Virtual Environment Workstation Project is to connect at least two of the current prototype interface systems to a common virtual environment database. The two users will participate and interact in a shared virtual environment but each will view it from their relative, spatially disparate viewpoint. The objective is to provide a collaborative workspace in which remotely located participants can virtually interact with some of the nuances of face-to-face meetings while also having access to their personal dataspace facility. This could enable valuable interaction between scientists collaborating from different locations across the country or even between astronauts on a space station and research labs on Earth. With full body tracking capability, it will also be possible for each user to be represented in this space by a life size virtual representation of themselves in whatever form they choose — a kind of electronic persona. For interactive theater or interactive fantasy applications, these virtual forms might range from fantasy figures to inanimate objects, or different figures to different people. Eventually, telecommunication networks will develop that will be configured with virtual environment servers for users to dial into remotely in order to interact with other virtually present users.

Although the current prototype of the Virtual Environment Workstation has been developed primarily to be used as a laboratory facility, the components have been designed to be easily replicable for relatively low-cost. As the processing power and graphics frame rate on microcomputers quickly increases, portable, personal virtual environment systems will also become available. The possibilities of virtual realities, it appears, are limitless as the possibilities of reality. It provides a human interface that disappears — a doorway to other worlds.



Figure 5

Notes:

1. Lipton, L. "Sensorama," *Popular Photography*, July, 1964.
2. Perkins, D. N. "Pictures and the Real Thing," Project Zero, Harvard University, Cambridge, Massachusetts, (1973).
3. Bender, *Environmental Design Primer*, Minneapolis (1973)
4. Lipmann, Andrew, "MovieMaps: An Application of the Optical Videodisc to Computer Graphics", *Computer Graphics* 14, #3 (1980).
5. Comeau, C. & Bryan, J. "Headsight Television System Provides Remote Surveillance", *Electronics*, (Nov. 10, 1961). Pp. 86-90.

6. Sutherland, I. E. "Head-Mounted Three-Dimensional Display", Proceedings of the Fall Joint Computer Conference, vol. 33. (1968), pp. 757-764.
7. Fisher, S.S. "Telepresence Master Glove Controller for Dexterous Robotic End-Effectors: Advances in Intelligent Robotics Systems, D. P. Casasent (ed) Proc. SPIE 726, (1986).
8. Fisher, S. S., McGreevy., Humphries, J., Robinett, W., "Virtual Environment Display System", ACM Workshop on 3D Interactive Graphics, Chapel Hill, North Carolina, (October 23-24, 1986).
9. Fisher, S. S., Wenzel, E. M., Coler, C., McGreevy, M. W., "Virtual Interface Environment Workstations", Proceedings of the Human Factors Society — 32<sup>nd</sup> Annual Meeting, Anaheim, California (October 24-28, 1988).
10. Wenzel, E. M., Wightman, F. L., Foster, S. H., "A Virtual Display System for Conveying Three-Dimensional Acoustic Information"" Proceedings of the Human Factors Society — 32<sup>nd</sup> Annual Meeting, Anaheim, California (October 24-28, 1988).
11. Foley, James, D. "Interfaces for Advanced Computing", Scientific American, 257, no. 4 (1987) pp. 126-135.