Research Directions in Virtual Environments

Report of an NSF Invitational Workshop

March 23-24, 1992

University of North Carolina at Chapel Hill

Gary Bishop, UNC-Chapel Hill (co-chair)
William Bricken, U. of Washington, Seattle
Frederick Brooks, Jr., UNC-Chapel Hill
Marcus Brown, U. of Alabama, Tuscaloosa
Chris Burbeck, UNC-Chapel Hill
Nat Durlach, M. I. T.
Steve Ellis, NASA-Ames Research Center
Henry Fuchs, UNC-Chapel Hill (co-chair)
Mark Green, U. of Alberta, Canada
James Lackner, Brandeis University
Michael McNeill, NCSA
Michael Moshell, U. of Central Florida
Randy Pausch, U. of Virginia, Charlottesville
Warren Robinett, UNC-Chapel Hill
Mandayam Srinivasan, M. I. T.
Ivan Sutherland, Sun Microsystems
Dick Urban, DARPA
Elizabeth Wenzel, NASA-Ames Research Center

TABLE OF CONTENTS

Executive Summary

At the request of NSF's Interactive Systems Program, a two-day invitational workshop was held March 23-24, 1992 at UNC Chapel Hill to identify and recommend future research directions in the area of "virtual environments" (VE) * . Workshop participants included some 18 experts (plus 4 NSF officials) from universities, industry, and other leading technical organizations. The two-day schedule alternated between sessions of the entire group and sessions in the following specialty areas, around which the recommendations came to be organized: 1) Perception , 2) Human-Machine Software Interface , 3) Software, 4) Hardware, and 5) Applications. Also, two participants developed a taxonomy of VE applications that is included as an appendix to the report.

Recommendations Summary:

Perception:

Vision

1. Collaborative science-technology development programs should be established at several sites around the country to encourage closer collaboration between developers and scientists.
2. Theoretical research should focus on development of metrics of performance and task demands in VE.
3. Paradigmatic applications and theoretical questions that illustrate the science-technology synergy need identification.

Audition

Spatial Sound

1. Theoretical research should emphasize the role of individual differences in Head-Related Transfer Functions (HRTF's), critical cues for distance and externalization, spectral cues for enhancing elevation and disambiguating the cone-of-confusion, head-motion, and intersensory interaction and adaptation in the accurate perception of virtual acoustic sources. The notion of artificially enhanced localization cues is also a promising area.
2. A fruitful area for joint basic and applied research is the development of perceptually-viable methods of simplifying the synthesis technique to maximize the efficiency of algorithms for complex room modeling.
3. Future effort should still be devoted to developing more realistic models of acoustic environments with implementation on more powerful hardware platforms.

Nonspeech Audio

1. Theoretical research should focus on lower-level sensory and higher-level cognitive determinants of acoustic perceptual organization, with particular emphasis on how acoustic parameters interact to determine the identification, segregation, and localization of multiple, simultaneous sources.
2. Technology development should focus on hardware and software systems specifically aimed at real-time generation and control for acoustic information display.

Haptics

1. Development should be encouraged of a variety of computer-controlled mechanical devices for either basic scientific investigation of the human haptic system or to serve as haptic interfaces for virtual environments and teleoperation.
2. Research programs should be initiated to encourage collaboration among engineers who are capable of building high precision robotic devices and scientists who can conduct biomechanical and perceptual experiments with the devices.
3. Research programs should also be developed to enable collaboration among researchers working on visual, auditory, and haptic interfaces, together with computer specialists who can develop software capable of synchronized handling of all the sensory and motor modalities.

Motion Sickness in Virtual Environments

---

*By *virtual environments*, we mean real-time interactive graphics with three-dimensional models, when combined with a display technology that gives the user immersion in the model world and direct manipulation. Such research has proceeded under many labels: *virtual reality, synthetic experience,* . etc. We prefer virtual environments for accuracy of description and truth in advertising.

1. The virtual environment community should be made aware of the sensory-motor adaptation and motion sickness problems to be expected presently because of hardware limitations and in the future as better virtual presence in nauseogenic environments is achieved.
2. Research programs should be initiated to evaluate the incidence and severity of sickness associated with different types of virtual environments, and to assess the kinds of sensory-motor adaptations and aftereffects associated with virtual environments.

Evaluation of Virtual Environments

Research should be conducted on the development of psychophysical techniques that measure the level of effort required to achieve a given level of performance, that relate performance on simple tasks with performance in a multi-task situation, and that operate in a systematic and well-defined manner with complex stimulus contexts.

Human-Computer Software Interface:

1. Researchers should focus on the development of new metaphors for VEs and the identification of reusable, application-independent interface components, specifically those which can be encapsulated in software and distributed.
2. NSF should support a software clearinghouse for code sharing, reuse, and software capitalization.
3. We will need to develop metrics to guide the exploration of VE tools, techniques, and metaphors.

Software:

1. The development of new modeling tools for model construction for virtual environments should be supported, especially inside-the-environment modeling tools. These tools need to be developed to the point where their effectiveness can be evaluated..
2. A facility for sharing existing and new models should be established.

Hardware:

Tracking Systems

1. Inertial tracking systems are prime for research activity now because of recent advances in micro-accelerometers and gyros. Inertial adjuncts to other tracking methods for sensing of motion derivatives is also a needed research activity.
2. Research into tracking technologies that allow large working volumes in outside spaces should be encouraged.

Haptic Systems

1. Support basic biomechanical and psycho-physical research on human haptic senses.
2. Support development of interactive force reflecting devices, and devices to distribute forces spatially and temporally within each of the (possibly multiple) contact regions.

Image Generators

1. Research into low latency rendering architectures should be encouraged.
2. Research is needed.into software techniques for motion prediction to overcome inherent system latencies and the errors they produce in registered see-through applications.

Visual Display Devices

NSF should primarily support pilot projects that offer potential for order of magnitude improvement in resolution, brightness and speed. NSF should also investigate display techniques that may offer decreases in latency and to characterize problems with display phenomena such as frame sequential color.

Applications:

1. Applications are needed which provide discriminatory power to evaluate VE technology versus ‘through the window’ interactive graphics and other similar technologies.
2. Researchers should look toward applications which solve real-world problems. VE must move beyond the stage of an interesting technological toy and begin to solve problems for people where they are.
3. Researchers should begin work on the probable impact of VE technology on society: Will VEs change the way we work (telecommuting/teleconferencing) or our interpersonal interactions? As the technology becomes more readily available, how will society react?

4. Can the use of VEs to communicate between people approach the level of communication we currently experience in person or in a group? What research must be done to move toward that goal? Is it even a desirable goal?

I. Introduction

What is Virtual Environments Research? In 1965, Ivan Sutherland in a paper, "The Ultimate Display", given at the triennial conference of the International Federation of Information Processing Societies, proclaimed a program of research in computer graphics which has challenged and guided the field ever since. One must look at the display screen, he said, as a window through which one beholds a *virtual world*. The challenge to computer graphics is to make the picture in the window look real, sound real, and the objects act real. Indeed, in the ultimate display, one will not look at that world through a window, but will be immersed in it, will change viewpoint by natural motions of head and body, and will interact directly and naturally with the objects in the world, hearing and feeling them, as well as seeing them.

Real-time interactive graphics with three-dimensional models, when combined with a display technology that gives the user immersion in the model world and direct manipulation, we call *virtual environments*. Such research has proceeded under many labels: *virtual reality, synthetic experience,* .etc. We prefer virtual environments for accuracy of description and truth in advertising. Merriam-Webster's New Collegiate Dictionary, Ninth Edition, defines

virtual as "being in effect but not in actual fact", and

environment as "the conditions, circumstances, and influences surrounding and affecting an organism".

Why is VE Research hot now? From 1965 until the mid-1980's, the limited power of computers and of graphical engines meant that Sutherland's vision could only be realized for crude depictions or for painfully slow interactions for many worlds. Many graphics researchers worked on making more faithful visual depictions by solving the problems of perspective, hiding, raster-scanning pictures, shading, or illumination. They got fidelity of motion by animation onto film, computing minutes per frame, giving up interaction. Others worked on real-time motions and interactions in toy worlds of only a few hundred elements.

Advances in technology, in computer and graphics organization, in displays, and in interactive devices now enable us to do in a video frame tasks that used to require batch computing. Digital signal processing algorithms and hardware allow the realistic production of three-dimensional sound cues, and increasingly compact and high performance mechanical sensors and actuators promise realistic simulation of manual interactions with objects. So it is now possible to bring these lines of research together and to approximate Sutherland's vision of interestingly complex worlds with rather good pictures, sounds, and forces, with tantalizingly close to real-time performance.

Though we still have far to go to achieve "The Ultimate Display", we have sufficiently advanced towards the goal that is timely to consider real systems for useful applications:

- What are the characteristics of the applications that will most benefit from such man-machine systems?

- What are the technical barriers that stand in the way of these applications?

- How can these most profitably be addressed? How can NSF (or DARPA) and the VE research community make a coordinated push through these barriers?

II. Overview

In light of the recent surge of interest in Virtual Environments in science, industry, and the media, an invitational workshop was held at the University of North Carolina at Chapel Hill on March 23-24, 1992, at the request of Dr. John Hestenes (Director, Interactive Systems, National Science Foundation). The workshop was chaired by Drs.. Gary Bishop and Henry Fuchs with the purpose of developing recommendations for research directions in this field. Eighteen researchers from the US and Canada spent two days in large and small groups developing a consensus on the recommendations in this report.

The participants divided into five working groups in order to focus on:

1. Perception (chaired by Steve Ellis),
2. Human-Computer Software Interface (chaired by Randy Pausch),
3. Software (chaired by Mark Green),
4. Hardware (chaired by Michael Moshell),and
5. Applications (chaired by Marcus Brown).

Also, two participants, Ivan Sutherland and Warren Robinett, developed a taxonomy of VE applications that is included as an appendix.

The recommendations of each of the groups were reviewed and discussed by all of the participants.

This report summarizes the results of the workshop. These results are organized around the five divisions of the working groups. Each section presents the current status of the sub-area, the perceived needs, and recommendations for future research directions.

III. Perception

Vision

Because of the pervasive, dominant role of vision in human affairs, visual stimuli are without question the most important component in the creation of the computer-based illusion that users are in a virtual environment. There are four aspects of this key role of vision: the characteristics of the visual image, the structure of the visual scene, the visual consequences of manipulative and vehicular interaction with the scene, and the role of visual information for spatial orientation.

Status

<u>Visual image</u> Modern visual psychophysics makes intensive use of computer graphics to synthesize high resolution stimuli for experimental manipulation. Display generation and digital filtering techniques have come to play an essential role in modern laboratories studying human vision. The mathematical and computational techniques used to describe the visual stimuli that are studied have also become the languages in which theories about visual phenomena are phrased (Watson, 1989)

<u>Visual scene</u> Structure in the visual image is automatically identified by biological image processing that segregates foreground from background and spontaneously groups regions together into subparts. Some aspects of this image segregation appear to be the result of parallel processing while other show evidence of sequential processing (Treisman, 1985). Once segregated, the contours and features collected into groups may be interpreted as objects in the space surrounding the observer. The separated patterns of contours and regions may then be interpreted as a surrounding space.

<u>Visual world</u> The spatial interpretation of visual images is highly dependent upon the kinematic characteristics of the image motion, in particular those motions that are consequences of the observer himself (Cutting, 1986). The patterns of image motion that are associated with observers' movements provide much of the necessary information for guidance through a cluttered environment and have provided

the basis for development of what J. J. Gibson described as a higher-order psychophysics. In this field, researchers may investigate the natural linkages established between properties of image, or object motion, and complex normal behaviors such a walking or object avoidance.

Just as motion of an observer causes global changes in the pattern of relative motion in the visual image, so to manipulative interaction with visible objects also produces characteristic visible transformations related to the object's position and identity, (e.g. Warren, et al, 1991), which have been extensively studied to provide the bases for psychological and physiological theories of manipulative interaction.

Visual orientation Visual information is not only important for local navigation while traversing an environment but also for global path planning and route selection. These more global tasks have been studied in isolation during scientifically motivated experiments (e.g. in Howard, 1982). But visual orientation is also important for more integrated tasks in which subjects use visual aids such as maps to maintain their internal representation of the surrounding space and assist planning of future activities.

Needs

Visual image Precision visual tasks will require improvements in the image quality of small display systems that provide photopic luminance levels with several arc-minute pixel resolution. Low level visual performance should be assessed with visual parameters likely to be provided by future display systems which may use nonstandard pixel layouts, variable field resolution, and field magnification to optimize allocation of computer graphics processing. Higher resolution inserts in the central visual field may be utilized, but gaze directed control of these fields may not be necessary if they can be made sufficiently large, i.e. to about 30 degrees. Since the presentation of wide fields of view (> 60 degrees monocular), will likely involve some geometric image distortion, studies of the tolerable distortion and characteristics of adaptation will also likely be required for specific tasks. However, because the binocular overlap between the left and right eye images need not be complete, monocular fields exceeding 60! may only rarely be required.

Visual scene Since virtual environment will only be able to present somewhat degraded low level visual cues such as contrast and stereopsis, the capacity for viewers to segregate foreground from background is likely to be less than that with natural images from real environments. Accordingly, visual segregation with degraded image quality and dynamics should be studied and enhancements to overcome difficulties should be developed.

Visual consequences The visual consequences of environmental interactions generally involve intersensory integration and do not quality as strictly visual issues. However there are purely visual consequences of motion in a simulation which are important for perceptual fidelity: a compelling visual simulation will require dynamic as well as kinematic modeling which currently is difficult to carry out at the necessary interactive rates, which ideally should exceed 30 Hz simulation loop frequency. Important work is required on the subjective and objective operator reactions to approximated kinematic and dynamic models of synthetic environments. How far can a simulation deviate from correct dynamical modeling and still appear to be realistic?

Visual orientation Imperfect and slow dynamics of virtual environments can lead to significant difficulties for users to maintain their spatial orientation within a simulated larger environment. Orientation aids to compensate for these difficulties should be developed to allow developers to simulate highly detailed real environments when such detailed simulation is required. These aids amount to enhancements for orienteering within a virtual environment and should assist users in switching between ego- and exocentric frames of reference which will be needed for efficient interpretation and control of objects in the simulated environment.

Recommendations

1. Collaborative science-technology development programs should to be established at several sites around the country to encourage closer collaboration between developers and scientists.

2. Theoretical research should focus on development of metrics of performance and task demands in VE.

3. Paradigmatic applications and theoretical questions that illustrate the science-technology synergy need identification.

Comment: The inherent interdisciplinary nature of VE will benefit from curriculum modifications to improve communication between perceptual scientists and interface designers. Currently, these researchers have significantly different research agenda and goals which can interfere with collaboration. Interface designers are happy with informal, imperfect guidance not the relative truth which the scientists seek.

Audition

Status

Two general areas of acoustic research, spatial sound and the real-time generation of nonspeech audio cues, are critical for virtual environment research and technology development. Speech generation and recognition, also important features of auditory displays, will not be discussed here.

Spatial Sound The simulation of spatial localization cues for interactive, virtual acoustic displays has received the most attention in recent work. Perceptual research suggests that synthesis of purely anechoic signals can result in perceptual errors, in particular, increases in front-back reversals, decreased elevation accuracy, and failures of externalization. These errors tend to be exacerbated when virtual sources are generated from non-personalized Head-Related Transfer Functions, a common circumstance for most virtual displays. In general, the synthesis technique involves the digital generation of stimuli using Head-Related Transfer Functions (HRTFs) measured in the ear canals of individual subjects or artificial heads for a large number of real source (loudspeakers) locations (e.g., Wightman & Kistler, 1989; Wenzel, 1992). Other research suggests that such errors may be mitigated by providing more complex acoustic cues derived from reverberant environments (Begault, 1991). Recently, some progress has been made in interactively synthesizing complex acoustic cues using a real-time implementation of the image model (Foster, et al., 1991)

Nonspeech Audio Following from Gibson's ecological approach to perception, one can conceive of the audible world as a collection of acoustic "objects". In addition to spatial location, various acoustic features such as temporal onsets and offsets, timbre, pitch, intensity, and rhythm, can specify the identities of the objects and convey meaning about discrete events or ongoing actions in the world and their relationships to one another.

One can systematically manipulate these features, effectively creating an auditory symbology which operates on a continuum from "literal" everyday sounds to a completely abstract mapping of statistical data into sound parameters. Principles for design and synthesis can be gleaned from the fields of music (Blattner, Sumikawa, and Greenberg, 1989), psychoacoustics (Patterson, 1982), and higher-level cognitive studies of the acoustical determinants of perceptual organization (Bregman, 1990; Buxton, Gaver, and Bly, 1989). Recently, a few studies have also been concerned with methods for directly characterizing and modeling environmental sounds such as walking sounds (Li, Logan, and Pastore, 1991). Other relevant research includes physically or structurally-based acoustic models of sound source characteristics such as radiation patterns (Morse and Ingard, 1968).

Needs

Spatial Sound It seems clear that simple anechoic simulations of spatial cues will not be sufficient to minimize perceptual errors and maximize perceptual "presence". Dynamic modeling of complex acoustic environments requires enormous computational resources for real-time implementation in a truly interactive (head-tracked) display. Currently it is not practical to render more than the first one or two reflections from a very small number of reflecting surfaces in real-time. However, because of the less stringent requirements of the auditory modality, acoustic digital signal processing is now advanced enough to allow significant strides in our basic understanding of human sound localization. While fully realistic, interactive simulations of a concert hall may not yet be feasible, synthesis techniques are sufficiently developed to allow an unprecedented degree of stimulus control for the purposes of psychophysical studies.

<u>Nonspeech Audio</u> A few cue-generation systems have been specifically integrated for virtual environment applications while some designers are beginning to develop systems intended for data "sonification" However, far more effort should be devoted to the development of sound-generation technology specifically aimed at information display. Perhaps more critical is the need for further research into lower-level sensory and higher-level cognitive determinants of acoustic perceptual organization, since these results will serve to guide technology development. Further, relatively little research has been concerned with how various acoustic parameters interact to determine the identification, segregation, and localization of multiple, simultaneous sources. Understanding of such interaction effects will be critical in any acoustic display developed for both virtual environments and telepresence.

Recommendations

<u>Spatial Sound</u>

1. Theoretical research should emphasize the role of individual differences in HRTFs, critical cues for distance and externalization, spectral cues for enhancing elevation and disambiguating the cone-of-confusion, head-motion, and intersensory interaction and adaptation in the accurate perception of virtual acoustic sources (see Wenzel, 1992). The notion of super-auditory localization, or artificially enhanced localization cues, is also a promising area (Durlach, 1991).

2. A fruitful area for joint basic and applied research is the development of perceptually-viable methods of simplifying the synthesis technique with the goal of maximizing the efficiency of algorithms for complex room modeling (increasing the number and complexity of modeled reflections).

3. In contrast to visual display technology, we are currently much closer to developing truly realistic simulations of auditory environments. Since the research pay-off is likely to be both high and timely, future effort should still be devoted to developing more realistic models of acoustic environments with implementation on more powerful hardware platforms. Some of the issues that need to be addressed are nonuniform radiators, diffuse reflections, scattering reflectors, diffraction and partial obscuration by walls or other objects, spreading loss and high-frequency absorption.

<u>Nonspeech Audio</u>

1. Theoretical research should focus on lower-level sensory and higher-level cognitive determinants of acoustic perceptual organization, with particular emphasis on how acoustic parameters interact to determine the identification, segregation, and localization of multiple, simultaneous sources.

2. Technology development should focus on hardware and software systems specifically aimed at real-time generation and control for acoustic information display, using basic theoretical knowledge as design guidelines.

Haptics

The human haptic system is composed of subsystems that enable tactile and kinesthetic senses as well as motor actions. In contrast to the purely sensory nature of vision and audition, only the haptic system is capable of direct action on real or virtual environments. Being able to touch, feel, and manipulate objects in the environment, in addition to seeing (and/or hearing) them, gives a sense of compelling immersion in the environment that is otherwise not possible. It is quite likely that much greater immersion can be achieved by the synchronous operation of even a simple haptic interface with a visual display, than by large improvements in the fidelity of the visual display alone. Consequently, it is important to develop a wide variety of haptic interfaces to interact with virtual environments. Examples of haptic interfaces that are being used in virtual environment research are joysticks and hand/arm exoskeletons. In general, they measure and display users' body part positions as well as the forces on them. The biomechanical, sensorimotor, and cognitive abilities of the human haptic system determine the design specifications for the hardware and software of haptic interfaces.

Status

Compared to vision and audition, our understanding of haptics is very limited. The recent trend of conducting multidisciplinary haptic studies involving biomechanical, psychophysical and neurophysiological experiments together with computational models has contributed to rapid progress. Due to the availability of powerful computers and high precision mechanical sensors and actuators, it is now possible to exert control over experimental variables as never before.

<u>Biomechanics of Contact:</u> In any task involving physical contact with an object, be it for exploration or manipulation, the mechanics of the contact interface plays a fundamental role. Both in the tactile sensory information flow from the object, and in imposing desired motor action on the object, the contact interface is strongly influenced by the surface and volumetric physical properties of the skin and subcutaneous tissues. Although some data on the in vivo biomechanical properties and several simple computational models of, say, the primate fingerpad are available, they are inadequate at present.

<u>Sensing and Control of Interface Variables:</u> The term interface variables is meant to include the kinematic variables (i.e., the relative positions, orientations, and motions) of various body parts, together with the associated normal and shear forces arising from contact with objects. The kinematic information is conveyed by our kinesthetic sense, whereas the contact forces are sensed by both tactile and kinesthetic systems. In the kinesthetic space, psychophysical phenomena such as anisotropies in the perception of distance and orientation, apparent curvature of straight lines, non-Euclidean distance measures between two points etc., have been reported. In tasks involving physical contact, such as manipulation of objects, it is known that tactile information is crucial in controlling grasp forces. These control actions can range from a fast spinal reflex to a relatively slow, deliberate action. However, even simple questions concerning our abilities, (such as what is our resolution in the sensing and control of interface variables), or the mechanisms, (such as how we perceive joint angles or contact forces), do not yet have unequivocal answers.

<u>Perception of Contact Conditions and Object Properties:</u> The contact conditions perceived through the tactual sense can be broadly classified as static (with or without skin stretch due to shear forces), slipping and vibratory. The object properties that are inferred include both geometric (such as shape), and material properties (such as compliance). The perception of both contact conditions and object properties is based on intensive, temporal, spatial or spatio-temporal stimulus variations, and the associated neural codes. Recent psychophysical and neurophysiological investigations have provided some answers to questions concerning perception and neural coding of roughness, raised features on rigid objects, slip, microtexture, shape, compliance, etc. However, the important connection between the loads imposed on the skin surface within the regions of contact with objects and the corresponding perception has only begun to be addressed.

Needs

In order to enunciate the specifications for the design of haptic interfaces, performance of the human haptic system should be characterized. This includes the determination of (a) the biomechanical properties of skin and subcutaneous soft tissues that govern the mechanics of contact with the interface; (b) the abilities of the human haptic system and the strategies used by human subjects in performing haptic tasks; and (c) evaluation of the effectiveness of haptic interfaces. A major barrier to progress from the perspectives of biomechanics, neuroscience and psychophysics has been the lack of robotic stimulators capable of delivering stimuli under sufficiently precise motion and force control.

<u>Biomechanical Investigations:</u> The tight mechanical coupling between the human skin and haptic interfaces strongly influences the effectiveness of the interface. Therefore, the specifications for the design of sensors and actuators in the interface, as well as the control algorithms that drive the interface, require the determination of surface and bulk properties of, say, the fingerpad. The measurement of force distributions within the contact regions with real objects is needed to determine how a display should be driven to simulate such contacts in virtual environments. In addition, computational models of the mechanical behavior of soft tissues will aid in simulating the dynamics of task performance for testing control algorithms, as well as in determining the required task-specific force distributions for the displays. This requires measurement of the in vivo skin and subcutaneous soft tissue response to time-varying normal and tangential loads.

Psychophysical Investigations: Determination of the human abilities (in terms of resolution, bandwidth, etc.) in sensing and control of net contact force vectors as well as joint angles or end point positions will set the design specifications for haptic interface devices. The rather large body of data available on tactile sensing of vibratory stimuli and the data on spatial localization and resolution, together with additional psychophysical experiments on the perception of contact conditions and object properties will influence directly the design of tactile displays. Theoretical developments concerning the task-specific flow of sensory information and control of motor action are needed to generate testable hypotheses on our haptic interactions with both real and virtual environments.

Haptic Interface Evaluation: Although the haptic interfaces available at present are quite limited in their capabilities, they need to be evaluated from the perspective of human perception. For example, a force-reflecting joystick attached to the floor can be called a grounded display, whereas a force-reflecting exo-skeletal device attached to the user's forearm would be an ungrounded display. (It would, in fact, be grounded at the forearm). The grounding choice affects whether or not the user experiences throughout his/her entire body the stresses induced by contact with a virtual object. The consequences of using an ungrounded display to simulate contact forces which really stem from grounded sources are not known and warrant investigation. Furthermore, the fidelity with which the tactual images have to be displayed and the motor actions have to be sensed by the interface depends on the task, stimulation of other sensory modalities, and interaction between the modalities. Experimenting with the available haptic interfaces, in conjunction with visual and auditory interfaces, helps to identify the necessary design improvements.

Recommendations

1. Development of a variety of computer-controlled mechanical devices for either basic scientific investigation of the human haptic system or to serve as haptic interfaces for virtual environments and teleoperation should be encouraged. The biomechanical and psychophysical research work detailed in the *Needs* section above should be supported.

2. Research programs should be initiated to encourage collaboration among engineers who are capable of building high precision robotic devices and scientists who can conduct biomechanical and perceptual experiments with the devices.

3. Research programs should also be developed to enable collaboration among researchers working on visual, auditory, and haptic interfaces, together with computer specialists who can develop software capable of synchronized handling of all the sensory and motor modalities.

Motion Sickness in Virtual Environments

Status

Human movement and locomotion involve a dynamic sensory-motor adaptation to the 1G background force of Earth. Disruptions of this relationship, which depends on correlations among patterns of sensory feedback from vision, touch, somatosensation, proprioception, and the semicircular canals and otolith organs of the labyrinth with motor information about ongoing movements, leads to a variety of perceptual and motor errors and often motion sickness as well. Errors and sickness persist until adaptation to the new situation is achieved. Different types of rearrangements differ in terms of how difficult it is to adapt to them and for some situations, adaptation cannot be achieved. There also are great individual differences in terms of susceptibility to motion sickness, ability to adapt, rate of adaptation and retention of adaptation.

Rearrangements can be of many types, including for example, delays in feedback loops as encountered in flight simulators or teleoperator systems, various types of visual or auditory position displacements, alterations of the effective inertial mass of the head or body such as are brought about by wearing headgear or a space suit, changes in the motor commands necessary to achieve movement such as in rotating environments or altered gravitoinertial force fields. Artificial transport of the body in vehicles - cars, trains, boats, aircraft, spaceships - leads to complex alterations in sensory-motor control which typically elicit performance errors and motion sickness until adaptation is achieved. Interestingly, sickness also develops during experienced motion of the body when the patterns of feedback present during voluntary movements are appropriate for the actual rather than experienced motion. This highlights the fact that no general theory exists which predicts which situations will be disruptive, nauseogenic, or difficult to adapt to, or which individuals will be most prone to these problems.

It can be expected, however, that sickness will be a problem both in the early development stages of virtual environments and once they have been perfected. In fact, reports of motion sickness are becoming commonplace. It can be fully expected that as sickness brought about by low resolution displays, visual lags, and improper force feedback disappears because of technical improvements, sickness will be more frequent because of the improved ability to create virtual environments involving "vehicle" motions which are naturally nauseogenic.

Needs

Motion sickness is going to be a highly significant problem in virtual environments. Moreover, it must be recognized that motion sickness is a complex syndrome. Nausea and vomiting represent only one part of motion sickness and often not the most crucial aspect of sickness. Other symptoms include headache, dizziness, eye strain, lethargy, and fatigue. Often a distinction is made between "gut" and "head" symptoms.

Under laboratory conditions, motion sickness is relatively easy to recognize. Experimental studies of motion sickness typically involve highly provocative test situations which rarely last more than an hour or two. Thus, sickness is expected and the personnel carrying out the experiments are highly skilled in recognizing signs and symptoms of acute motion sickness. Often the subjects are trained in identifying the subjective concomitants of sickness as well. Motion sickness is more difficult to recognize under operational conditions. Such conditions tend to be less provocative and to bring on initially more head symptoms than gut ones. A sailor may not realize that his experience of drowsiness is an early sign of developing motion sickness. Highly motivated individuals may be able to "tough out" head symptoms in order to complete a work schedule, but ultimately the generation of work schedules will be constrained.

An attempt should be made to familiarize investigators in the virtual environment area with the primary characteristics of motion sickness. In addition, it would be useful for records to be kept concerning the incidence and characteristics of sickness encountered with different test platforms. Other characteristics of the users, such as motion sickness histories, should also be gathered. In all likelihood, research programs involving the experimental evaluation of sickness with different types of dynamic virtual environments will be necessary as well as programs to study ways of enhancing the rate of acquiring adaptation and of enhancing retention of adaptation. Relevant here, too, will be the assessment of the consequences of adapting to virtual environments with regard to sensory-motor performance on return to the normal environment. We anticipate that functionally significant disruptions of performance will be associated with adaptation to dynamic virtual environments. Their severity, as well as ways of eliminating them, need to be addressed.

Recommendations

1. The virtual environment community should be made aware of the sensory-motor adaptation and motion sickness problems to be expected presently because of hardware limitations and in the future as better virtual presence in nauseogenic environments is achieved.

2. Research programs should be initiated to evaluate the incidence and severity of sickness associated with different types of virtual environments.

3. Research programs should be developed to assess the kinds of sensory-motor adaptations and aftereffects associated with virtual environments.

Virtual Environments for Perception Research

Status

Scientists using computer-generated stimuli and computer-based signal processing have been using the elements of virtual environment technology for years as part of the normal course of research. Computer processing, especially graphics, is regularly used to synthesize sensory images and present experimental stimuli. Though the most obvious presentations have involved visual stimuli, auditory, haptic, and vestibular stimuli are also commonly presented. Typically, these experimental applications are not focused on synthesis of multisensory environments. They are, however, necessarily carefully calibrated and often automated for the conduct of an experiment. Digital signal processing is also extensively used to filter and analyze performance data (See Elkind, Card, and Hochberg, 1989).

Virtual environment (VE) technology is inherently an interdisciplinary field that will be likely to integrate the previous scientific work in perception and manual control (Ellis, 1991; Durlach et al, 1992). There is consequently a great potential benefit from collaboration between perceptual/motor researchers and interface developers. VE can produce physically unrealizable stimulus environments that are uniquely useful for testing theories of perception and manual control, but good/standard calibration techniques are needed to avoid unwitting investigation of display artifacts.

Needs

The natural course of this research will require the development of appropriate evaluation metrics and calibration procedures to be suggested for use by VE system developers. Though anecdotes from successful users indicate significant benefits from the technology, existing VE systems need to be generally more analytically evaluated for specific utility.

VE development forces evaluation of assumptions/better design of tests and can force psychological and physiological theories to be more precisely stated and integrated and can enable the asking of totally new experimental questions.

Evaluation of Virtual Environments

Status and Needs

VE systems can be meaningfully evaluated only by determining their effects on the human user. No other measure matters. Rigorous, reliable and interpretable measures of these effects require careful and exact control of the stimulus and use of well-designed psychophysical procedures. Although there is a strong well-established tradition of sensory research that conforms to these requirements, evaluation of VE systems presents demands that cannot all be met by existing techniques. New procedures must be developed.

Many of the well-developed experimental techniques focus on the detection of small differences in the stimulus, e.g. in contrast detection, spatial discriminations, motion detection and discrimination and depth discriminations. Among the psychophysical methods currently in use, those that use high contrast stimuli are most likely to be generally useful. Motion detection and velocity discrimination, for example, may prove to be useful indicators of the perceptual stability of the system. Similarly, size discrimination may prove to be a useful indicator of the quality of the depth percept.

Also of interest is the veridicality of the simulated percept. Subjective judgments of absolute size or depth may be useful, but rating judgments are highly context-dependent. They may miss substantial differences between real and virtual environments because each is judged only relative to itself. See-through head-mounted displays that permit combination of virtual and real environments offer a means of determining the veridicality of the simulation by direct comparison to the real environment.

However useful these techniques, none addresses directly the question of how demanding the system is to use. Psychophysical experiments are marked by a strong restriction of task and stimuli. Virtual environments, on the other hand, typically present complex stimuli to observers who make complex judgments. Breaking these complexities down into simple components that are amenable to traditional psychophysical evaluation cannot answer many of the most important questions about the value of the virtual environment because performance depends on how information is acquired in the face of the complexity.

On the other hand, use of complex stimuli can make interpretation of the data difficult. The observer may use stimulus cues other than the one being manipulated intentionally. This problem can be avoided by manipulating only a single localized feature and letting the rest of the stimulus serve as context. In that case, a large number of different contexts or environments must be used to prevent the observer from learning to ignore the context. Use of multiple varied contexts inevitably adds noise to the measured response, requiring more trials, which requires more contexts, etc. Soon the number of images to be stored becomes large enough to be problematical. Restricting the contexts to a class that is appropriate to a single application reduces the variability at the cost of generality. Techniques need to be developed for systematically generating and describing contexts to be used in well controlled studies. A generally accessible library of contexts might add order to this pursuit.

At least as important as stimulus complexity is the problem of task complexity. The accuracy with which an observer can perform a single task tells us little about the resources required to support that performance. For virtual environments to be useful, they cannot place unreasonable demands on the user. If it is difficult or taxing for the observer to collect each desired piece of information, the system may not enhance performance of complex tasks. The information presented in the VE must not only be available, it must be easily available.

Most psychophysical methods do not deal with this aspect of the problem. Some that do exploit the temporal dimension. Reaction times may be useful indicators of the difficulty of a task. This measure has the notable advantage that the task being studied need not be limited to detecting barely discriminable stimulus differences. The major disadvantage to this technique is the need to handle speed-accuracy tradeoffs. Another approach is to limit the exposure time of the stimulus. This avoids the problem of speed-accuracy tradeoffs, and can result in a more systematic representation of the speed of processing. The major weakness of the approach is that for all but contrast detection tasks, a masking stimulus must be presented at termination of the stimulus to stop further processing. Choice of an effective mask is often difficult and because one cannot know whether a given mask is completely effective for a given stimulus, comparison of effects across stimuli is never wholly convincing. Improvements in these techniques and/or creation of new techniques that make use of the temporal dimension to assess task difficulty would be helpful.

More generally, we need to develop ways to assess behavior when the observer is performing more complex tasks, e.g., when he is performing several simple tasks simultaneously or in rapid alternation. Some techniques that have been used in other areas of experimental psychology, e.g., in memory research, may be usefully adopted here.

Recommendation

Research should be conducted on the development of psychophysical techniques that measure the level of effort required to achieve a given level of performance, that relate performance on simple tasks with performance in a multi-task situation, and that operate in a systematic and well-defined manner with complex stimulus contexts.

IV. Human-Computer Software Interface

Status

The interaction techniques for virtual environments have yet to be extensively explored. There are several design principles which are likely candidates for structuring VE interface metaphors, such as:

- natural behavior as an interaction principle
- rapid prototyping for design space exploration

- •knowledge-based agents for interactive guidance
- •respect for the physiological context of the physical body in VE design
- •supernormal capabilities as a metaphor for sensory remapping

These principles are suggestive, but need to be evaluated in task specific application domains.

Current difficulties in the design and construction of virtual worlds include:

- •ineffective software, particularly for programming dynamics and interactivity;
- •complex hardware configurations and especially hardware latency;
- •world modeling and maintenance (version control on worlds is complex);
- •no theory of world construction;
- •physical literalism (assuming the virtual world is like the physical);
- •naive models of interaction in a virtual environment; and
- •failure of intuition (physical solutions are not virtual solutions).

Current research challenges include:

- •developing software operating systems which facilitate effective, inclusive, real-time interaction, multiple participant, multiple sensory, high bandwidth, low latency VEs;
- •determining which aspects of the 2D WIMP and desktop metaphors are generalizable to higher dimensional virtual worlds;
- •identifying the core of generic virtual world design tools and application independent interaction techniques;
- •integration of multidisciplinary teams to address human, software, and hardware interactions, including design and analysis of experiments; and
- •bringing multidisciplinary knowledge to the construction of VEs, including work from database, data fusion, networking, human factors, computer supported collaborative work, and artificial intelligence communities.

Needs

The most pressing needs are:

- •software for design of and interaction with virtual worlds that is modular, flexible, and abstract, particularly interpreted languages for rapid prototyping;
- •software operating systems and infrastructure to support world design, construction, and interaction, particularly software which reduces latency;
- •metaphors which guide the exploration and prototyping of particular tools and techniques for use in VEs;
- •measurement techniques and theories for identifying differential effects of world designs on the sense of presence;
- •measurement techniques for identifying resource expenditure, cognitive load, transfer of learning, adaptation effects, and other performance parameters of different configuration of VEs; and
- •task-specific evaluation of software tools.

Secondary, less general, needs include the development of:

- •navigation techniques for high dimensional data and displays;
- •location and movement techniques and software tools;
- •manipulation techniques and software tools;
- •event history, filtering, and recording tools;
- •behavioral and dynamic models for determining the dispositions of objects;
- •specification languages for world dynamics;
- •editing tools for objects and for environmental spaces, including models of inter-object communication, process management, and composition rules;
- •a design theory of sensory presentation modes (which sensory suites are best for conveying which tasks?);

- •languages and grammars for describing world interactions;
- •the virtual body, mapping tools for connecting sensors on the physical body to an accurate physiological model in software, and in turn, to the virtual object being used as a body in the VE; and
- •tools that can be used for interaction and construction both inside the VE and outside on a monitor viewing the VE.

Recommendations

Since VE design and interaction is in its infancy, these recommendations are focused on generic rather than specific goals.

1. Isolating and evaluating application-independent interaction techniques and metaphors. Researchers should focus on the development of new metaphors for VEs and the identification of reusable, application-independent interface components, specifically those which can be encapsulated in software and distributed. One specific area with high potential is the use of voice input as a parallel input modality.

   While some of this evaluation will be extensive research centered on the human's capabilities, some of it will be rapid, less formal evaluation to help interface designers choose between conflicting alternatives. In one sense, these different objectives underscore the differences between basic science and engineering. We explicitly suggest that NSF recognize the contributions made by evaluations made at various levels of certainty. Fred Brooks refers to these as *certainty shells*, including findings, observations, and rules-of-thumb. This research needs to integrate the diverse skills and styles of multidisciplinary teams.

2. Software capitalization — NSF should support a software clearinghouse for code sharing, reuse, and software capitalization. The cost of having each VE laboratory develop its own infrastructure is prohibitive to the effective conduct of research. We encourage support of world building and maintenance tools, to enable version control, composition of components developed by different design groups (tool portability), ease of customization and configurability, and expressability.

3. Measurement techniques to determine the quality of VEs We will need to develop metrics to guide the exploration of VE tools, techniques, and metaphors. The quality of a VE is likely to be related to specific tasks, physiological comfort, cognitive and perceptual load, resource expenditure, intuitiveness of tools and interactions, ease of reconfiguration, transfer and adaptation effects, and even individual differences and preferences. In this regard, our final, concrete recommendation is that NSF help forge inter-disciplinary collaborations, especially between computer scientists and psychologists.

V. Software

Status

The cost effective development of virtual environments will require a new generation of software tools. These tools must support the special hardware devices that are used in virtual environments and the design of the environments themselves.

A number of software tools have been developed to support the production of virtual environments. These tools can be divided into two broad groups based on whether they are commercial products or the result of a research project.

Most of the commercial products support a particular hardware configuration that is sold by the software vendor. They provide a basic level of software support for the vendor's hardware and allow for the development of simple virtual environments. Most of these tools provide basic support for the hardware devices and limited support for the development of the virtual environments themselves.

The research software tools tend to be more general, since there is no mandate to support a particular hardware platform. These tools support a range of hardware configurations, which facilitates the sharing of research results and the resulting program code. Again, most of these tools only address the device support issues, but in a hardware independent manner. Some tools are beginning to be developed for supporting the production of virtual environments, in addition to supporting the basic hardware. One of the major advantages of this group of software tools is that they allow new researchers to quickly enter the field in a cost effective manner, since they are freely distributed and will likely support whatever hardware configuration the researcher has available.

Needs

Modeling. Modeling is the key software issue at the present time. The modeling problem is not unique to virtual environments, but it is crucial to the success of many VE applications. In many VE applications, there is a tight connection between the user and the underlying model. The user must be able to move through the model, interact with the objects in the model, and change the model interactively. The development of good modeling tools will facilitate the development of new applications and the development of new techniques.

In this report, modeling refers to the data structures that are used to record the geometrical information for the environment. This information includes the shape of the objects in the environment, their moving parts and physical properties, and the behaviors that they can perform (how they interact with other objects in the environment and with the user). This information is not only used to produce the visual presentation of the environment, but it is also used in sound production and tactile feedback.

In a VE application there may be multiple users for the same application. The users could have different needs or be performing different tasks, which would result in a different view of the model for each user. In addition, each user may have personal preferences that dictate a different view of the model. The modeling software most be able to support the different views of the users.

There are also applications where multiple users may be in the same environment at the same time. This implies that the model must be shared among the users. Each user must have a consistent view of the application (which may depend upon how the model is viewed). There must be methods for keeping these models up-to-date and at the same time allow the users to modify individual copies of the model.

The model will be used to drive several different media such as graphics and sound. The model must contain the information required by all of these media, plus the information required to synchronize their presentation. This is not as simple as events in two media occurring at the same time. For example, in the case of lightning, the visual appearance of the lightning must occur before its sound is produced. This problem also occurs in multi-media systems.

Objects in a VE may correspond to objects that occur in the real world. When this happens, we expect these objects to behave in the same way as real world objects behave. In other words, they must follow the traditional laws of physics. When an object is dropped, it should fall until it reaches the ground. It shouldn't be possible for objects to pass through each other. In order to get these effects we need to be able to model at least part of physics. This includes the familiar laws of motion and the ability to detect collisions between objects. These techniques must perform in real-time, so they can be incorporated into the feedback provided by the environment.

Some of the objects in the VE must be able to respond to the user's actions and the actions of other objects. There should be ways of specifying this behavior and assigning these behaviors to the objects in the model. Some of the objects in the environment may also have autonomous behaviors, such as walking. This will involve interactions between the objects and simulated time within the environment. There is some overlap between the techniques required here and the techniques that are being developed in computer animation and simulation.

One of the issues in modeling for VE is whether the modeling should be done inside or outside of the environment. There are clear advantages to both approaches. Modeling in the environment gives the designer a good view of what the user will see when he or she is in the environment. Modeling outside the environment allows the designer to use multiple views of the environment and draw upon the modeling techniques that have been developed over the past few decades.

Current inside-the-environment modeling systems are not as good as the outside-the-environment ones. There are two main reasons for this. First, the I/O devices used in VE are not as well developed as the devices used in traditional interactive computer graphics. For example, head-mounted displays have a much lower resolution than CRT monitors, and most 3D tracking technologies have higher noise levels than 2D devices. Second, the design techniques used in the outside-the-environment modelers have been developed over the past few decades, while there has been very little experience with inside-the-environment modelers. There should be an effort to develop design techniques for inside-the-environment modelers.

One of the main problems with large models is navigating in them. In 2D applications, overview maps of the model can be presented in separate screen windows, but similar techniques haven't been developed for 3D environments. It has been suggested that the environment can be scaled in order to give an overview of it. This will only work if detail can be culled from the model when it is scaled. In order to do this effectively, the model must be constructed in such a way that it is easy to identify the details and the important components of the model. The model should provide assistance to the navigation process.

Modeling is currently a time-consuming process. We need better tools to support this task, particularly modeling tools for VE applications. Reasonable modeling tools have been developed for other domains, most notably computer-aided design and product design. The objects in both of these domains are static, and thus these tools don't address the complete range of VE modeling requirements. An effort should be made to develop better modeling tools that address the problems of VE modeling. The development of these tools will assist in other domains that make heavy use of modeling.

Since the production of good models currently requires a considerable amount of effort, there should be some mechanism for sharing the existing models. There are two aspects to this sharing. First is the determination of a standard format for encoding and transmitting the models. There are numerous formats that have been developed for the interchange of graphical information. One of these formats may be suitable for this purpose, or a new format may need to be developed. The format that is used should be able to encode all the modeling information, including hierarchical structure, physical properties of the objects, sound and behavior. Second, there should be one or more sites that store, maintain and distribute these models. These sites would be responsible for collecting the existing models and new models as they are developed, cataloguing the set of models and distributing them to researchers on request. All of this should be performed over the existing computer networks to guarantee the widest distribution of the models.

There is a need for certain real-time functionality in VE systems. In particular, there is a need for low latency in the processing of requests, the ability to assign priorities to processes, and the management of time. Real-time operating systems exist that meet these requirements, but researchers haven't been willing to use them, since they would need to give up the facilities that are provided by existing main-line operating systems, such as UNIX. As a result, we would like to see these real-time facilities added to UNIX.

Our recommendation is not to directly fund this work under the VE initiative but to encourage other groups to fund this effort under larger projects that are also be interested in this technology. In particular, the high performance computing initiative should be encouraged to fund this work. Other areas that would benefit from this work include robotics and real-time simulation.

Recommendations

1. The development of new modeling tools that meet the requirements of model construction for virtual environments should be supported.

2. A facility for sharing existing and new models should be established. This will involve both the development of standards for model interchange, and the establishment of one or more sites where the models will be maintained.

3. Support the development of inside-the-environment modeling tools. These tools need to be developed to the point where their effectiveness can be evaluated.

VI. Hardware

Hardware for virtual environments is an area of rapid progress, for several reasons. Much of the technology has other applications outside of VE, in areas such as robotics, entertainment and instrumentation. Areas such as high definition display devices are being driven by the general advancement of semiconductor design and fabrication.

Nevertheless, a substantial number of "show stoppers" for VE can be found in hardware. For instance, without some breakthroughs in the bio-mechanical engineering of haptic (force and touch feedback) devices, many of the envisioned applications of VE will remain science fiction.

We will explore the following hardware categories: tracking systems, haptic systems, image generators, visual display devices, audio systems and speech input systems.

Tracking Systems:
This section concerns techniques for determining the location and orientation of the user's head, hands and ultimately of any real-world object of interest to the VE user.

*Status***.**

Five categories of devices exist:

1. Magnetic systems, typified by the Polhemus and Ascension trackers, use magnetic fields emitted by small antennas and detected by one or more receiving antennas. They are unreliable in the presence of metal objects. Today's systems have a range of less than 1m from the emitter; multiple-emitter solutions are possible.

   Latencies (time from initiation of a motion until the new data arrives at the host computer via a parallel link) in current systems are on the order of 50 msec. Rumored improvements may bring this down to 5 msec within a year. Current systems can achieve resolutions of approximately 1mm and 0.03 degrees, and accuracies of approximately 3mm and 0.1 degrees.A wireless version of a magnetic tracker would be technically difficult.

2. Acoustic systems, typified by the Logitech Mouse, use ultrasonic pulses. Range and latency are approximately the same as today's magnetic devices, without susceptibility to magnetic interference. They are susceptible to line-of-sight occlusion, e. g. by the arm of the user. Wireless versions are feasible. Ultimate latency is limited by the speed of sound (about 0.3 m/msec). Accuracy is limited by variations of the speed of sound with changes in air density.

3. Inertial systems have formerly been bulky, expensive, and had too much drift. Advances in micromachines show promise of producing small, inexpensive accelerometers and rate gyros that have sufficient sensitivity to allow tracking via dead reckoning for short periods. Wireless operation would be feasible. These devices are not yet commercially available.

4. Among mechanical systems, the BOOM viewer from Fake Space Labs uses a rigid framework both to support the viewing device and measure position and orientation, with a reported accuracy of 4mm and 0.1 degree resolution at each joint. Because, unlike the magnetic and acoustical devices, no averaging is required, the latency is very small - under 1 msec. Low-cost, non-load-bearing mechanical trackers are also available.

5. Optical systems such as the Optotrack 3000 and the Selspot II determine position of targets via triangulation techniques from cameras at known locations. They determine orientation by observing multiple targets. Without large separations between targets, fine measurement of orientation is very

difficult. These systems provide only a small working volume (1 meter cube) and are susceptible to line of sight occlusion.

Optical systems such as the experimental tracker at UNC mount the cameras on the head and observe targets at fixed locations on the ceiling. This system determines both position and orientation directly and may be extended to cover arbitrary areas but must maintain line of sight and isn't directly suitable for tracking the hands.

Needs.

To support most goals currently under consideration, a position sensor would need to have these characteristics:

- wireless operation
- ability to track multiple users in same space without interference
- range of up to 10m x 10m x 5m with reference to a base unit, perhaps the size of a briefcase;
- no wide-area antenna, ceiling or sensor field required
- latency under 5 msec
- resolution of 1 mm and 0.01 degree
- accuracy of 1 mm and .01 degree in registered see-through applications, 1 cm and 0.1 degree in non-registered applications.
- sampling rate >= 60 Hz
- direct sensing of state AND derivatives

The private sector is making progress towards better tracking systems for VE though the technology hasn't changed significantly for more than 10 years.

Recommendations.

1. Inertial tracking systems are prime for research activity now because of recent advances in micro-accelerometers and gyros.

2. Inertial adjuncts to other tracking methods for sensing of motion derivatives is also a needed research activity.

3. Research into tracking technologies that allow large working volumes in outside spaces should be encouraged.

Haptic Systems:

This section concerns devices which provide the illusion of manual exploration or manipulation of objects in virtual worlds. These devices employ human tactile, kinesthetic and motor systems for interaction with the virtual environment. They perform some or all of two basic functions:

1. measurement of the positions and forces (and time derivatives) of the user's hand and/or other body parts to be used as control inputs to VE, and

2. display of forces and positions and/or their spatial and temporal distributions to the user.

*Status***.** Four areas of work are of interest:

1. Hand position/joint angle measurement: Instrumented wands, joysticks, gloves and exoskeletons that measure hand position/joint angles are available in the market. The major problems are the intrusion the user feels while wearing, say, an exoskeleton, and the ever-present need for improvements in resolutions and ranges.

2. Application of forces and torques: Exoskeletons attempt to both measure hand/arm joint angles and to load these joints with feedback torques around the joint by applying forces at the contact regions. With

inputs from research into the teleoperation of robots, some progress has been made on the exoskeletal problem, but the systems are heavy, expensive, clumsy and unreliable. The computational problem of compensating for the mass of the control arm is substantial.

Simulator motion bases represent the most mature component of this sub-area. The art of providing sustained forces to the subject's whole body through a pilot's seat is well understood. The principal limitations are that the average acceleration provided by a terrestrial motion base must always be at least 1g, and that sustained accelerations greater than 1g are achievable only in centrifuges, which induce other severe problems such as coriolis effects.

3. Tactile Displays. These are two dimensional fields of force applied to the skin, to simulate touching. Both normal and shear forces are necessary to simulate general sensations of touch. The simulation of different coefficients of friction (for various materials and skin conditions: wet, dry, oily) is problematic. A few standalone demonstrations of tactile displays have been built.

4. Other stimulus distributions. These include thermal, several kinds of pain, and possible direct electrical stimuli applied as two dimensional fields to the skin. Pain is the report of a physiological limit's being reached or exceeded. "Deliberate synaesthesia" would use a small vibrating stimulator or small shock (rather than pain) to report that a "virtual limb" or teleoperation actuator device has reached its limit.

Needs.

Mechanical stability is important. Force feedback systems need to have vibration rigorously controlled, probably to less than a few microns amplitude, to prevent false cues. Forces on the order of 10 Newtons are needed, with 10 bit resolution. With respect to force distributions, a spatial density exceeding 1 mm/taxel and a temporal resolution approaching 1 kHz are needed.

Recommendations.

1. Support basic biomechanical and psycho-physical research on human haptic senses.

2. Support development of interactive force reflecting devices, and devices to distribute forces spatially and temporally within each of the (possibly multiple) contact regions.

Image Generators

Status.

The computer graphics industry is rapidly improving its polygon, lighting models and texture rendering capability, and the cost of visual systems continues to drop at a rate >50%/yr for constant performance.. The principal deficit for VE applications concerns latency - the delay between a change to the visual database or viewing parameters and a change in the display.

The latency of fast workstations such as Silicon Graphics products is one frame interval (33ms for 30Hz update rate) to compute the image in the offscreen buffer, plus one refresh interval (14ms at 72Hz refresh). The drive in the private sector towards higher through-put will not automatically solve the latency problem, as pipelined architectures will probably continue to be used. The VE community needs lower latency even at the expense of polygon through-put.

Needs.

Most graphics users do not need low latency (at the levels required by VE), and so most graphics architectures are deeply pipelined. Research in computer architecture is needed to explore ways to reduce latency.

There is no good data on the threshold for perception of latency; research is needed. Simulator sickness studies on the effect of latency have generally explored the domain between 100 and 300 msec, because of limitations in the imaging systems in use.

One way to determine this threshold would be with the use of a panoramic buffer. A pre-stored scene would be "panned" in response to head tracking information. The inherent latency of an NTSC display device is around 16 ms; in order to beat that rate, "racing the beam" would be necessary (i. e. the generation of scan lines immediately before their rendering, based on most current geometry).

The lowest latency VE systems likely to be built in the near future will still have delays from user motion to visual output that cause significant errors in registered see-through systems and perceptual/sickness effects in all VE systems.

Recommendations.

1. Research into low latency rendering architectures should be encouraged.
2. Research into software techniques for motion prediction to overcome inherent system latencies and the errors they produce in registered see-through applications is needed.

Visual Display Devices

Status.

Commercially available LCD displays are presently in the range of 200 x 300 pixels (consisting of a color triad). DARPA is funding development of a true 640 x 512 electroluminescent, and also a color LCD system consisting of color quads (1280 x 1024 in monochrome). The displays are to be ready within 2 years.

DARPA is also funding the development of a reflective system using very small deformable mirrors; around 10**6 moving mirrors. The device will have high speed (10 microsec) and power capabilities because all energy is reflected, not absorbed. Can be used with color by triplex or field seq color.

Optics for head mounted displays are available, but are heavy and expensive. Because the images in VE are totally synthetic, it is possible to perform some transformations (such as correction of some kinds of chromatic aberration) in software which formerly required optical methods.

Needs.

1024 x 1024, color, 60 Hz or faster, in 1" square packaging; lightweight, large visual field optics.

Recommendation

1. NSF should primarily support pilot projects that offer potential for order of magnitude improvement in resolution, brightness and speed.

2. NSF should also investigate display techniques that may offer decreases in latency and to characterize problems with display phenomena such as frame sequential color.

Audio Systems:

Status

Spatial Sound Briefly, the approach is to synthesize externalized, three-dimensional sound cues over headphones using very high-speed digital signal processing devices (see Wenzel, 1992). In general, the synthesis technique involves the digital generation of stimuli using Head-Related Transfer Functions (HRTFs) measured in the ear canals of individual subjects or artificial heads for a large number of real source (loudspeakers) locations (e.g., Wightman & Kistler, 1989). In most current systems (e.g., the

Convolvotron), from one to four moving or static sources can be simulated (with varying degrees of fidelity) in a head-stable, anechoic environment by convolution of arbitrary signals with HRTF-based filters chosen according to the output of a head-tracking device. Motion trajectories and static locations at greater resolutions than the empirical data are generally simulated by linear interpolation between the measured impulse responses. Also, in some systems, a simple distance cue can be provided via real-time scaling of amplitude.

Nonspeech Audio The ideal synthesis device would be able to flexibly generate the entire continuum of nonspeech sounds described earlier as well as be able to continuously modulate various acoustic parameters associated with these sounds in real-time. Current devices available for generating nonspeech sounds are based almost exclusively on MIDI (Musical Instrument Digital Interface) technology and tend to fall into two general categories; "samplers", which digitally store sounds for later real-time playback, and "synthesizers", which rely on analog or digital sound generation techniques originally developed for imitating musical instruments. With samplers, many different sounds can be reproduced (nearly) exactly, but substantial effort and storage media are required for accurately pre-recording sounds and there is usually limited real-time control of acoustic parameters. Synthesizers are more flexible in the type of real-time control available but less general in terms of the variety of sound qualities that can be generated. A potential disadvantage of both is that they are not specifically designed for the generation and control of sounds for information display and tend to require that the user have specialized knowledge of musical/production techniques.

Needs

Any generally useful audio hardware system must sample its inputs with at least 16 bits of resolution and 40 to 50 khz sampling rate. To perform the spatial synthesis functions described in the Status section, a computational rate of about 300 MIPs is required. Additional computational power will probably be needed to implement more complex environmental models as well as the real-time generation and control of nonspeech audio cues.

Recommendations:

This seems to be an area where the technology has arrived at a price/performance level which can support rapid progress. Research should now be funded primarily to improve models of perception and techniques for generating acoustical models of the environment and of environmental sounds (acoustical "objects"), rather than for extensive hardware development. However, while not the primary emphasis, some new hardware development will probably be required to accommodate more elaborate and efficient models of complex environments as they develop.

Speech Input:

Status

Inexpensive devices are now available which, with reasonable reliability, can recognize individual words from the speaker who trained the system. Speaker-independent methods are achieving some acceptance.

Needs:

Many selection tasks would be better mediated by voice than by pointing, given the unsatisfactory state of visual pointing in 3D. As the technology of speech recognition improves, textual input in some situations could move from keyboard to voice, which will impact VE-based database activities and other information accessing tasks.

Recommendations:

Support would be appropriate for projects which explore the integration of voice input into the VE user interface. Hardware development is probably not needed.

VII. Applications

Status

For VE systems and technologies with respect to particular applications, there are two basic questions:

1. Can the given application be accomplished with VE technology? Many applications are currently beyond the state of the art.
2. If the application can be accomplished, will it prove superior to other technologies for accomplishing the same task?

While the promise of VE technology has been widely acknowledged, there are very few production-quality applications that are used regularly to solve real-world problems. Many desired applications have requirements which are currently beyond state of the art. For those that are feasible, it remains to be shown that VE provides a superior solution.

Compared to computer models displayed via conventional screen-bound interactive ("through-the-window") graphics, VE offers

- •total immersion in the computer model,
- •kinesthetic, intuitive viewpoint change, and
- •direct manipulation in 3-D.

Thus, one would expect to be able to show comparative superiority in simple applications that exploit these advantages.

These comparative advantages promise an order of magnitude better illusion for the user, thus greater involvement of the user with the task. This promise alone has been the driving force for development of VE technology in the past. With the exception of vehicle simulators and entertainment, most current applications of VE have been developed for driving and testing the technology rather than to accomplish ends related to the applications.

The following application areas stand to gain significantly from developments in VE technology:

1. Data or Model Visualization:

Scientific visualization techniques are widely accepted as a means for extracting understanding from large data spaces. When VE technologies are coupled with existing visualization systems, the user will experience a fundamental paradigm shift in interacting with the data or model. Stereoscopic imaging combined with intuitive control over point of view frees the user to concentrate on the data rather than the computer interface. The incorporation of other senses into the "display" potentially offers a mechanism for correlation of data features through non-visual means.

Augmentation is an extension of data visualization: The image presented to the user is a combination of a computer-generated image and the view of the real world around the user.

2. Designing and Planning:

A primary characteristic of design activities is the iterative *analyze-refine* cycle which takes place as the model evolves. Any method which aids the designer during this process will improve the entire activity. Because a strength of VE is the capability for direct manipulation of objects within the virtual space, design activities should benefit greatly from this technology. Existing applications for architectural design and surgical planning attest to the benefits of VE technology.

3. Education and Training:

Computer models of real-world systems can be presented through a virtual environment. Given that the appropriate level of fidelity can be provided, the user can interact with the virtual system as though it were the real thing. In this context, the educational implications of this technology are immediately obvious. In addition to simply portraying the real world, the rules governing a virtual world can be altered to allow the user to experience and learn from alternate "realities". In this way, the effect and importance of various physical parameters can be studied, and the educational focus can be on cognitive skills rather than the specifics of a particular environment.

4. Teleoperations:

VE can be applied to situations where an environment is too hazardous for a human, but a tele-sensor/operator may enter. As an example, the exploration of the sunken Titanic was accomplished primarily through teleoperations. Although the interface was not called 'VE,' this is an obvious extension. (Although it is an interesting problem; research into the robotic teleoperators is beyond the scope of VE research. VE research should focus on the interface between the human and the computer model, not on the model-world interface.)

5. Psychological Test Beds:

VEs can produce physically unrealizable stimulus environments that are uniquely useful for testing theories of human perception and manual control. This includes research into visual, auditory, and haptic sensory stimuli, and its effects, both short-term and long-term, on the user of a VE.

6. Entertainment and Artistic applications:

Given the amount of public and media interest in VE technology, there is an economic potential for entertainment which bills itself as "virtual reality". There are existing applications of this type currently installed in major theme-parks. One can only expect to see this market expand in the future as public exposure grows.

7. Artistic Applications

Artists are often the first individuals to explore new media and technologies. The basic activity of an artist is creating space — whether physical, mental, visual, or emotional. As VE technology becomes more widely accessible, it is reasonable to expect that artists will begin to explore the unique possibilities for expression that virtual environments offer.

8. Communication and Collaboration

Virtual environments coupled with high-speed networks and distributed computation provide a common neutral space where communication, negotiation, and collaboration can take place. The ability for two or more people in separate locations to meet inside a virtual space and work toward a common goal has the potential to revolutionize the way collaborative work is done.

Needs

To date, most of the major deficiencies of VE have been demonstrated by experimentation with various applications. These deficiencies are not specific to the applications but to VE technology. They include:

1. Hardware, software, and interface issues.
   These have been discussed in detail in an earlier portion of the report.

2. Model engineering.
   Many researchers have found that building a reasonably detailed model for a non-trivial object (i.e., a house) takes an inordinate amount of effort (perhaps a man-year). The need for improvement is obvious.

3. Psychological measurement techniques and metrics.

Measuring the possible superiority of VE technology often involves measuring human cognitive performance. This type of measurement is extremely difficult.

4. Simulator sickness.

Experience with space flight suggests that the more realistic VE becomes and the longer people spend immersed in a VE, the more people in VE's may experience various physical reactions, commonly known as simulator sickness. This phenomenon is also discussed in an earlier portion of the report.

Recommendations

1. Researchers in VE should look toward applications which promote measurement. NSF should encourage applications which provide discriminatory power to evaluate VE technology versus 'through the window' interactive graphics and other similar technologies.

2. Widespread development of VE applications are dependent on the availability of a stable and robust software infrastructure. NSF should address the issue of making such a software environment available as new developers enter the field. The alternative is to require each new site to waste valuable time "reinventing the wheel".

2. NSF funding should be directed toward applications which solve real-world problems. VE applications must move beyond simple demonstrations of technology and begin to solve problems for people where they are.

3. If VE applications are to be taken seriously, the causes and effects of sickness from virtual environments must be well understood. To this end, research into "simulator sickness" should be supported. VE also seems to be an excellent platform for related and unrelated psychological/physiological measurements.

4. Interpersonal Interaction as facilitated by VE. Can use of VEs to communicate between people approach the level of communication we currently experience in person, or in a group? NSF funding should support research toward resolving these questions.

5. Researchers should begin work on the probable impact of VE technology on society: Will VEs change the way we work (telecommuting/teleconferencing)? Will they modify our interpersonal interactions? As the technology becomes more readily available, how will society react?

References

Begault, D. R. (1991) Perceptual effects of synthetic reverberation on 3-D audio systems. 91st Convention of the Audio Engineering Society, New York, Preprint 3212 (W-6).

Blattner, M. M., Sumikawa, D. A., & Greenberg, R. M. (1989). Earcons and Icons: Their Structure and Common Design Principles. Human-Computer Interaction, 4, 11-44.

Bregman, A. S. (1990). Auditory Scene Analysis. Cambridge, MA: MIT Press.

Buxton, W., Gaver, W., & Bly, S. (1989). The use of non-speech audio at the interface. (Tutorial No. 10). Presented at CHI'89, ACM Conference on Human Factors in Computing Systems, New York: ACM Press.

Durlach, N. I. (1991) Auditory localization in teleoperator and virtual environment systems: Ideas, issues, and problems. *Perception*, *20*, 543-554.

Durlach, N.I., Pew, R.W., Aviles, W.A., DiZio, P.A., and Zeltzer, D.L. (Eds.) (1992). Virtual Environment Technology for Training (VETT). The Virtual Environment and Teleoperator Research Consortium (VETREC). Bolt, Beranek, and Newman Report 7661.

Cutting, James E. (1986) Perception with an eye for motion. Cambridge, MA: MIT Press.

Elkind, Jerome I., Card, Stuart K., Hochberg, Julian, and Huey, Beverly M. (1989) Human performance models for computer-aided engineering, Washington D.C. National Academy Press.

Ellis, Stephen R. (1991) The nature and origin of virtual environments: a bibliographical essay. *Computer Systems in Engineering, 2*, 4, 321-47.

Ellis, Stephen R., Kaiser, Mary K., and Grunwald, Arthur J.(eds) (1991). Pictorial communication in virtual and real environments, Taylor and Francis, London,and Bristol, NJ.

Foster, S.H., Wenzel, E.M., and Taylor, R.M. (1991) Real- time synthesis of complex acoustic environments . *Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics,* New Paltz, NY.

Howard, Ian (1982) Human visual orientation, Wiley, New York.

Lackner, J.R. Some aspects of sensory-motor control and adaptation in man. In: R.D. Walk and H.L. Pick, Jr. (Eds.). Intersensory Perception and Sensory Integration, New York, Plenum, 143-173, 1981.

Lackner, J.R. Human sensory-motor adaptation to the terrestrial force environment. In: D. Ingle, M. Jeannerod and D. Lee (Eds.). Brain Mechanisms and Spatial Vision., Nijhoff, Amsterdam, 175-210, 1985.*

Li, X., Logan, R. J., and Pastore, R. E. (1991) Perception of acoustic source characteristics. Journal of the Acoustical Society of America, 90, 3036-3049.

Morse, P. M. and Ingard, K. U. (1968) Theoretical Acoustics. New York: McGraw-Hill.

Patterson, R. R. (1982). Guidelines for Auditory Warning Systems on Civil Aircraft. (Paper No. 82017), London: Civil Aviation Authority.

Treisman, A. (1985) Preattentive processing in vision. Computer Vision, Graphics, and Image Processing, 31 156-177.

Warren William, H., Jr., Mestre, Daniel R., Blackwell, Arshavir W., and Morris, Michael W. (1991) Perception of circular heading from optical flow. Journal of Experimental Psychology, 17, 28-43.

Wenzel, E. M. (1992) Localization in virtual acoustic displays. Presence: Teleoperators and Virtual Environments, 1, 80-107.

Wightman, F. L. & Kistler, D. J. (1989) Headphone simulation of free-field listening I: stimulus synthesis. Journal of the Acoustical Society of America, 85, 858-867.

Appendix

Taxonomies for Virtual Environments

On Taxonomies In General Taxonomy can help thinking about a group of ideas. A taxonomy characterizes ideas along several more or less independent axes so as to form a matrix. The taxonomy has two important parts, the axes and the entries in the matrix.

The items listed on each axis are related by whatever characteristic is represented by that axis. If the variable chosen is continuous, the axis will represent a spectrum of values which may be more or less well defined. If the variable chosen has only a few fixed values, then those values divide the space of all ideas considered into categories. For example, a taxonomy of people might include eye color as a variable, characterizing eye color into a small number of common values, e.g. blue, brown, hazel, etc. The several axes of the taxonomy divide the intellectual space of interest into a number of boxes, each of which can be examined in turn.

The boxes in a taxonomy are used to enumerate ideas that have common characteristics in the axes chosen. One can list in each such box a set of ideas that are related to each other by their position on the axis. If all of the axes are continuous, the taxonomy describes a space of possibilities. In such a diagram, each related collection of ideas occupies a region of the overall space instead of a discrete box.

The number of boxes in a taxonomy is, of course, its volume, i.e. the product of the number of values permitted to each variable. The most useful taxonomies generally have fewer than 100 boxes. With more boxes it is difficult to think about each one separately. Some useful taxonomies have only a few boxes; as few as four are sometimes useful. The desire to limit the number of boxes limits the number of values permitted to each variable. Of course if there are more variables, each must be permitted fewer values. A useful discrete variable usually provides 2-5 values; a useful continuous variable generally approximates continuity with small, medium, and large.

There are two parts of building a taxonomy that make building it interesting. First there is the selection of axes and values along them. In forming the axes, one is forced to think of orthogonal dimensions. Which variables are related and which are independent. It often happens that as the taxonomy takes shape one finds that the variables chosen are not really independent. For example, in forming the taxonomy presented here we initially chose a separate axis to distinguish "sensed real world" versus "simulated world." As the taxonomy formed we considered that maybe "simulated" could be included as a single value along the "time" axis. We considered making this change because we found inadequate distinction between boxes differing in this dimension.

The second interesting part of building a taxonomy is filling in the boxes. Here the really interesting results come out by examining combinations of variables that may hitherto not have been considered together. For example, our taxonomy that plots different sense modalities against different physical phenomena leads one to ask what smells look like, how pictures sound, and other such ideas that might not otherwise occur to unaided thinking. Making a taxonomy is the only organized way I know to go about creative thought.

Our Taxonomies for Virtual Environments We actually made three taxonomies, which divide up the space of all possible virtual environments in different ways.

<u>Data Flow Between the World and Human.</u> The first taxonomy comes about by observing that a mechanical or electronic device can be interposed between a human being and the real world, as shown in Figure 1 below.

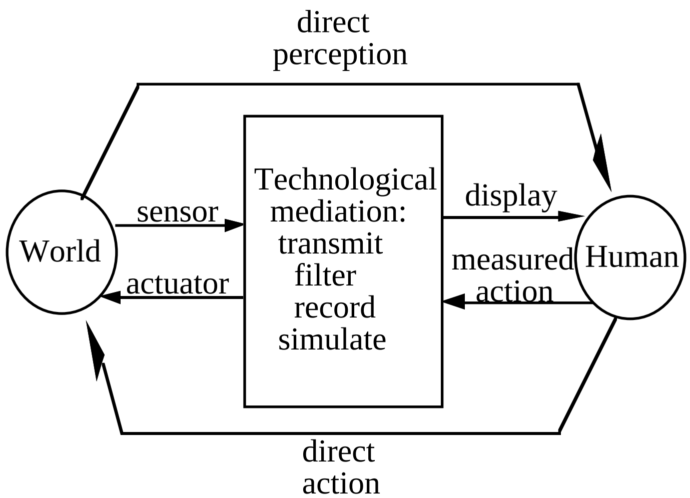

Figure 1. Technologically-mediated experience.

Inputs and outputs from the real world and the human are connected to this black box. This diagram covers all types of technologically-mediated experience, in which some sort of apparatus mediates human perception of and action within the world. The most important subcategories of technologically-mediated experience are recorded, transmitted, and simulated experience. Figure 2 shows that these types of mediated experience are distinguished by different patterns of data flow. The data flow for a robot is also shown in Figure 2, with possible supervision by a human operator. We include within the black box in each of the diagrams the *model* which defines the virtual world perceived by the human.

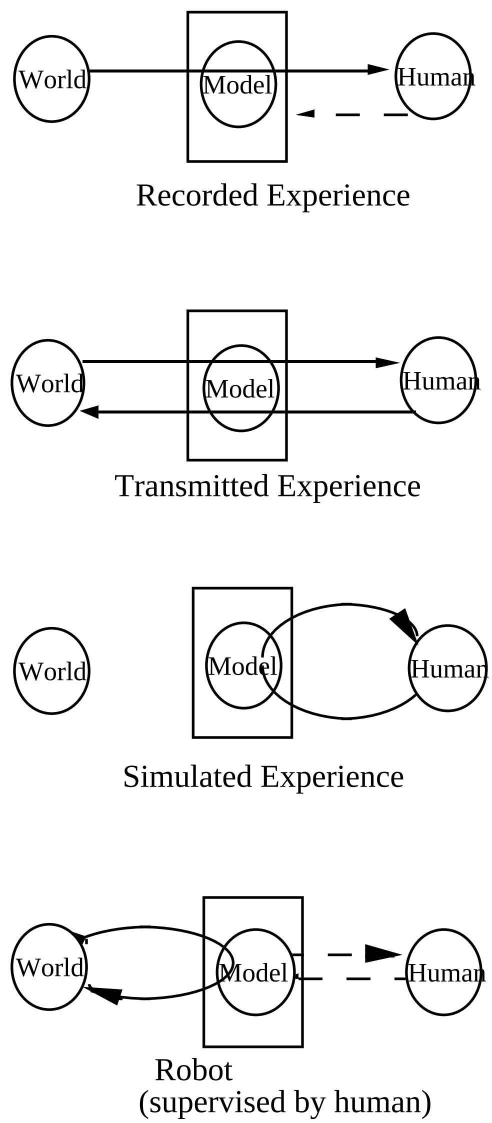

Figure 2. Data flow for types of mediatedexperience.

Two or more people may communicate or interact within a virtual environment, which leads to the data flow diagrams in Figure 3. The first two diagrams emphasize two different things, communication versus interaction with the virtual world. However all of the data paths exist to do both simultaneously. The last diagram of Figure 3 shows that two people may interact collaboratively with the real world through a shared representation of the real world.

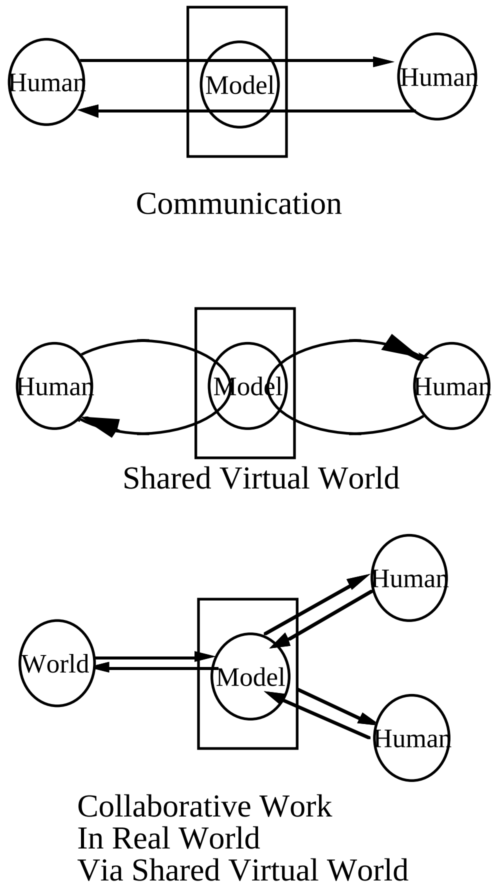

Figure 3. Two people in shared virtual world.

Time and Space Our second taxonomy chooses as axes three variables: time, space, and simulated vs real. In this taxonomy, the dimensions of time and space compare the time and space coordinates of the real world as captured by a sensor versus the actual physical time-space location of the human immersed in a virtual world. For example, in teleoperation, the space coordinates of the virtual world perceived by the human are different from the human's space coordinates.

We have characterized both time and space as "aligned", "displaced", and "scaled". By aligned we mean in real time or in spatial registry. By delayed we mean passed on through time by some kind of recording or displaced through space as in one of the many "tele" kinds of objects. By scaled we mean speeded up, slowed down, enlarged, or minified. Notice that if scaled, aligned does not apply unless two scale factors happen to match.

We have debated how to treat the "simulated versus real" dimension. One of us claims that "simulated" is really another value on the time dimension because in a simulation "real time" is a matter of choice. The other claims that simulations can be intended to model physical systems and therefore that "real time" does have a meaning within a simulation. Entering this debate is instructive in terms of generating ideas. For this report, we have left "simulated versus real" as a separate dimension of the taxonomy.

A issue that comes up in trying to classify simulated virtual worlds is that while some simulations are intended to model the real world, and thus can be classified as real-time or spatially registered, other simulations are completely fanciful and don't match up with any time or place in the real world. Thus "fantasy world simulation" seems to be a separate category of its own in our taxonomy.

We have drawn the taxonomy hierarchically in two stages. The top level is shown in Figure 4, and the two boxes "recording or transmission of real world" and "simulation of real world" are further broken down in Figure 5 along the dimensions of relative time and space. The entries in the boxes of Figure 5 are examples, not category definitions.

	Sense Real World	Simulate
Real World	recording or transmission of real world	simulation of real world
Fabricated World	?	simulation of fantasy world

Figure 4. Top level taxonomy of virtual environments.

Sensed Real World

	Transmit (real-time)	Record (1-to-1 time)	Record (scaled time)
Aligned (spatially)	night vision goggles	ghost (on-site replay of past events)	ghost speeded up
Displaced (spatially)	teleoperation	movie	slow motion instant replay
Scaled (spatially)	micro-teleoperation	micro-movie	micro-movie slowed down

Simulation of Real Worlds

	Record (1-to-1 time)	Record (scaled time)
real-time in-place sim	historical sim (reenactment on site)	historical sim speeded up
predictive sim in teleoperation	flight sim	sim of trip to Mars speeded up
real-time weather prediction	molecular sim	molecular sim slowed down

Figure 5. Taxonomy of time, space, and simulated vs. real.

Sensors versus Human Senses Our third taxonomy looks at extending human perception within and across sense modalities. Here one axis is just the sensory modalities. The other axis is various phenomena that could be sensed. This matrix is shown in the figure below. Neither axis has a complete list of senses nor phenomena.

	See	Hear	Feel	Vestibular	Smell
Visible Light	spectacles; night vision goggles		Opticon image-to-tactile transducer		
Sound	sonogram	hearing aid			
Force	visualization of strain in special plastic		teleoperation		
Inertial Changes	seismograph	rattle			
Chemical Compostion	gas chromatograph				smell amplifier
Ultrasound	medical ultrasound image	sonar			
Radio Waves	radio telescope	radio			
Infrared	night vision goggles				
X-rays	fluoroscope				
Magnetism	compass				
Radiation		Geiger counter			

Along the main diagonal of this matrix are augmentations and prostheses for the ordinary "built-in" human senses. For example, spectacles appear in the visible light row under vision, and hearing aids are listed at the intersection of sound and hearing. The diagonal entry for smell prompts the question "Could a device for amplifying smell be created?"

The off-diagonal boxes would contain examples of sensory substitution. For example, what do smells look like? What do sounds look like? How do sounds smell? It is the ability of this taxonomy to generate fresh ideas that is interesting.

The lower part of the matrix lists phenomena for which no corresponding human sensory system exists, such as ultrasound. For these imperceptible phenomena, mapping sensors which detect them to human senses expands human awareness by creating "synthetic senses."

Isolated versus Merged Another distinction not captured in any of the earlier taxonomies in this report is that a virtual world may either be portrayed to the human with the real world blacked out, or the virtual world may be superimposed onto the human's direct perception of the real world. This may be done either optically with half-silvered mirrors, or computationally by capturing a representation of the real world with cameras and then mixing the captured model of the real world with another model representing the virtual world.

It is most useful to merge the real and virtual worlds when the virtual world is spatially registered with the real world. However, even when the virtual world models a distant location, seeing through to the surrounding real world may be useful simply to let the human user avoid running into walls and tripping over obstacles.

Acknowledgements

The workshop from which this report proceeded was funded by National Science Foundation Contract Number
IRI-9213777 (Interactive Systems Program, Dr. John Hestenes, Director). We gratefully acknowledge the assistance of Ronald Azuma, James Chung, Stefan Gottschalk, Robert Katz, Mark Mine, Jannick Rolland, Shannon Stephenson, Kathy Tesh, Fay Ward, and William V. Wright, as staff assistants.