

VIRTUAL INTERFACE ENVIRONMENT WORKSTATIONS

S.S. Fisher, E.M. Wenzel, C. Coler, M.W. McGreevy
Aerospace Human Factors Research Division
NASA Ames Research Center
Moffett Field, California 94035

ABSTRACT

A head-mounted, wide-angle, stereoscopic display system controlled by operator position, voice and gesture has been developed at NASA's Ames Research Center for use as a multipurpose interface environment. This Virtual Interface Environment Workstation (VIEW) system provides a multisensory, interactive display environment in which a user can virtually explore a 360-degree synthesized or remotely sensed environment and can viscerally interact with its components. Primary applications of the system are in telerobotics, management of large-scale integrated information systems, and human factors research. System configuration, research scenarios, and research directions are described.

RESEARCH OBJECTIVES

The primary objective of this research is to develop a multipurpose, multimodal operator interface to facilitate natural interaction with complex operational tasks and to augment operator situational awareness of large-scale autonomous and semi-autonomous integrated systems. The system should specifically provide a uniform interface for multiple task supervision, be reconfigurable for varying levels of operator skill, training and preference, and have an operator interface configuration that features human matched displays and controls for transparent, natural system interaction and reduced training requirements.

In the Aerospace Human Factors Research Division of NASA's Ames Research Center, an interactive Virtual Interface Environment Workstation (VIEW) has been developed to aid in design, simulation and evaluation of advanced data display and management concepts for operator interface design. The VIEW system provides a virtual auditory and stereoscopic image surround that is responsive to inputs from the operator's position, voice and gestures. As a low-cost, multipurpose simulation device, this variable interface configuration allows an operator to virtually explore a 360-degree synthesized or remotely sensed environment and viscerally interact with its components.

Application areas of the virtual interface environment research are focused in:

1. Development of workstations for complex operational tasks such as telerobotic and telepresence control of remotely operated robotic devices and vehicles that require a sufficient quantity and quality of sensory feedback to approximate actual presence at the task site.

2. Development of concepts and guidelines for 'portable' multi-modal information management systems such as an EVA Spacesuit visor display, with subsequent development of workstations for supervision and management of large-scale integrated information systems in which data manipulation, storage and retrieval, and system monitoring tasks can be spatially organized.

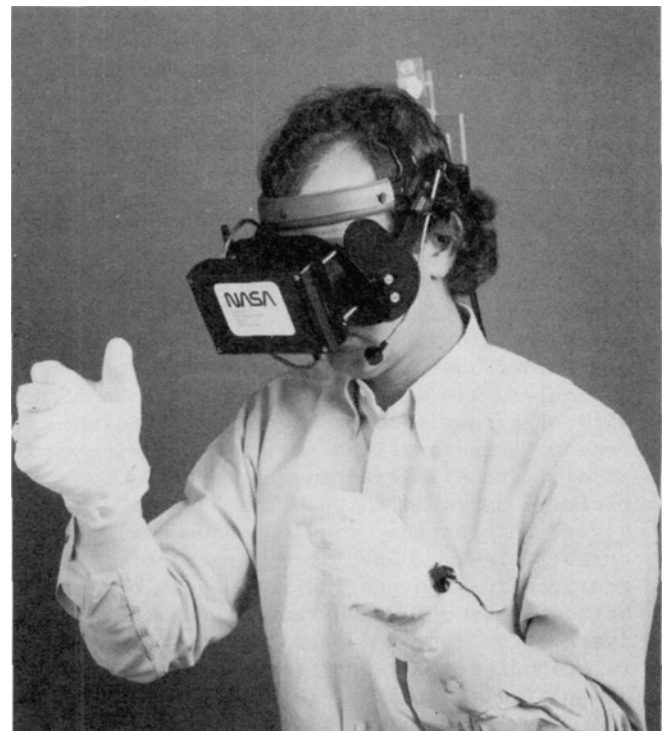


Figure 1. Head-coupled, stereoscopic display system with DataGlove tactile input technology for virtual object manipulation.

A more general research objective includes use of this display system to synthesize interactive test environments for aerospace human factors research in such areas as: spatial habitability research; rapid prototyping of display and workstation configurations; research on effective transfer of spatial information and visualization of multidimensional data; and spatial cognition research on multisensory integration.

SYSTEM CONFIGURATION

The current Virtual Interface Environment Workstation system consists of: a wide-angle stereoscopic display unit, glove-like devices for multiple degree-of-freedom tactile input, connected speech recognition technology, gesture tracking devices, 3D auditory display and speech-synthesis technology, and computer graphic and video image generation equipment. This hardware is integrated with a realtime Unix workstation that supports the computations required to drive an external high-performance realtime 3D graphics system, processes input from up to 12 realtime input peripherals (e.g., the trackers and gloves), and provides other realtime task processing. A collection of software called the 'simulation framework' has also been developed that consists of a well-documented library of functions to provide access to all of the system peripherals and software services, and of a collection of source files and simulation software that demonstrates the use of the major hardware and software components that make up the VIEW system in order to facilitate system reconfiguration for changing research requirements.

VIEW Visual Display Technology

The VIEW Display subsystem is an integrated technology package comprised of image display elements, optics, and electronics which together provide a wide angle, stereoscopic image environment that closely matches human binocular visual capabilities. When combined with high-resolution, magnetic 6 degree-of-freedom head and limb position tracking technology, the displayed imagery appears to completely surround the user in 3-space and provides interactive viewing and manipulation capabilities. Specific areas of display research in progress include: development of orthostereoscopic and remote presence display systems for improved operator performance in telerobotics and telepresence control; display of multi-dimensional, graphic representations of knowledge-based information systems to facilitate supervision and management of large-scale integrated systems; and 3D simulation of veridical experimental environments

for human factors research.

Head-mounted displays. The head-coupled display unit uses two passive, twisted nematic, monochromatic, liquid crystal display screens presented to each eye of the user through wide-angle optics. Resolution of each display is 640x220 pixel elements with 320 distinguishable vertical lines and approximately 16 levels of grayscale. The LCD elements incorporate a diamond-shaped pixel geometry that tends to reduce the apparent fixed pattern spatial noise by effectively doubling the horizontal and vertical fixed pattern spatial frequencies. Currently, the overall size of each display is 3.2" diagonally in a 4:3 aspect ratio. The transmissive LCD screens are backlighted with miniature fluorescent light sources that result in a contrast ratio of approximately 7:1. This LCD technology is capable of address times of 63.5 microseconds per line with typical pixel response times of approximately 32 milliseconds. Imagery displayed on the screens is a NTSC standard video signal generated by computer, remote video cameras, or a combination of these input media with other video sources such as optical video disk. Horizontal position of the images on the display screens can be electronically controlled by sliders on the display casing to adjust for image/optics alignment and for variable convergence requirements. The display unit is connected to support electronics that provide power and conditioned video signals through a 20 foot cable. In addition, these electronics provide stereo alignment patterns, user selection of biocular or binocular imagery output to the viewer, and can simultaneously output multiplexed stereo video signals for videotape storage and replay.

Optics. The display screens are viewed directly through a pair of wide angle magnifying lenses that are mounted approximately 5 mm from the screens with adjustments for variable accommodation requirements such as relative tilt and distance of LCD substrate from optics. These optics provide a 120 degree horizontal and vertical field of view for each eye and up to a 90 degree binocular field overlap. Total instantaneous visual field of view is approximately 120 degrees. The 2.75" diameter of each optical element requires a minimum 4" diagonal display size to completely fill the available field of view. With 640x220 pixel resolution, each pixel in the diagonal array (i.e. each distinguishable vertical line) subtends 0.38 degrees (22.5 minutes) of horizontal visual field. The wide angle optics create a pincushion distortion in the displayed imagery that requires a barrel distortion compensation in the image generation or image capture technology for correct

scene representation.

Mounting configurations. The display and optics unit is positioned directly in front of the viewer's eyes and coupled to head motion by means of a lightweight headgear configuration. Adjustments are provided for varying eye-relief distance and size of the supporting head band. Eye-cups around the optic elements are included to reduce interference of ambient light. The entire unit is counterweighted to transfer the weight of the display and optics unit to a point directly over the operator's spine. Also included on the headgear unit are: a support for attaching the head position tracking sensor, a microphone for input to the connected speech recognition subsystem, and earphones for auditory display feedback to the operator. See Figure 1.

Desk-mounted viewer. The wideangle viewing optics package is also implemented in a movable arm-mounted, workstation configuration for evaluation in design and engineering applications such as the display of three dimensional graphic databases used in Computational Fluid Dynamics Research. This viewer incorporates CRT display technology with a display resolution of 400 x 400 pixels. The position and orientation of the display unit and mounting arm is transduced and provided to the host computer to allow the operator to change viewpoints within the database in real time.

Image generation and capture. The display subsystem is used to view Virtual Environments that are synthesized with 3D computer-generated imagery, or that are remotely sensed by user-controlled, stereoscopic video camera configurations. The computer image system enables high performance, realtime 3D graphics presentation at resolutions of 640x480 and 1,000 x 1,000 pixels. This imagery is generated at rates up to 30 frames per second as required to update image viewpoints in coordination with head and limb motion. Dual independent, synchronized display channels are implemented to present disparate imagery to each eye of the viewer for true stereoscopic depth cues. For realtime video input of remote environments, two miniature CCD video cameras are used to provide stereoscopic imagery. Development and evaluation of several head-coupled, remote camera platform and gimbal prototypes is in progress to determine optimal hardware and control configurations for remotely controlled, free-flying or telerobot mounted camera systems. Research efforts also include the development of realtime signal processing technology to combine multiple video sources with computer generated imagery.

VIEW Auditory Display Technology

Binaural auditory display. The initial auditory subsystem for VIEW is capable of presenting a wide variety of binaural sounds to the operator via headphones using sound-synthesis technology developed for music synthesizers. This system has been developed so that a basic sound library may be designed and stored offline for later realtime manipulation using an interactive editing interface and/or 'front-panel' capabilities of the synthesizer. Thus, different sound signatures (e.g., different timbres or temporal patterns) can be associated with different objects or types of information in the visual displays to achieve a quasi-semantic or symbolic capability. Then, for each of the independent and simultaneous sound signatures or 'objects', a set of global sound parameters can be dynamically coordinated with the other display parameters via the interface between the host computer and the synthesizer. For example, stereo panning, pitch, and intensity can be associated with position, size, or distance of objects in the visual display. Also, commercially-available speech-synthesis technology with unlimited vocabulary is used to provide an additional display capability for voice report requirements and verbal acknowledgement of system input. This synthesizer is a text to speech converter that generates a human-sounding voice with capability to modify pronunciation of output words, to change speaking rate, and to select from six different voice qualities.

The primary function of this initial auditory display is to provide both discrete and dynamic auditory cues which can augment or supply information missing from the visual or gestural displays. For example, discrete cues might signal contact between telerobot end-effectors and target objects or alert the operator to attend to information in a data window which is currently out of the field-of-view. Similarly, auditory parameters can be dynamically modulated and coordinated with the other subsystem displays for such tasks as monitoring the relative positions of dual robot arms in order to keep them within a proscribed motion envelope.

3D Auditory display. Additional research is underway at Ames to develop an auditory display prototype that is capable of synthetically generating three-dimensional sound cues in realtime. These cues are presented via headphones and are perceived outside of the head at a discrete distance and direction in the 3-space surrounding the user. When integrated into the VIEW system, position of the operator can be monitored in realtime and the information used to maintain several

localized sound cues in fixed positions or in motion trajectories relative to the user. This capability will further aid the operator's situational awareness by augmenting spatial information from the visual display and providing outside the field-of-view navigational and cueing aids (see Wenzel, et.al., 1988, this volume).

VIEW Interaction

Visceral interaction with the Virtual Environment surrounding the user is enabled by speech and gesture input technologies. Current research includes configuration and evaluation of these input modalities, both separately and in combination, to facilitate effective, natural user interaction in various simulated task environments.

Speech recognition. The VIEW system includes commercially-available, speaker-dependent, connected speech-recognition technology that allows the user to give system commands in a natural, conversational format that can not be achieved with highly constrained discrete word recognition systems or through keyboard input. Typical speech mediated interactions are requests for display/report of system status, instructions for supervisory control tasks, and verbal commands to change interface mode or configuration.

DataGloves. For tactile interaction with the displayed three dimensional virtual environment, the user wears lightweight glove-like devices that transmit data-records of arm, hand and finger shape and position to a host computer. The gloves are instrumented with fiber optic flex-sensing devices at each finger joint and between fingers. Electromagnetic tracking sensors like those described for sensing the operator's head coordinates are also mounted on each glove to transmit relative position and orientation of the hands and arms to the host system. The system also includes preprocessor resident subroutine and editing capability to recognize and transmit specific hand/finger gestures for use in supervisory control modes, and software capability to quickly calibrate system measurements for use with different operator hand sizes. Primary uses of this technology are to provide a three-dimensional cursor in the displayed environment, and to enable interaction with virtual or remote objects. Current implementations of this research at Ames include a three-dimensional graphic database of an articulated hand that, in the virtual display environment, appears spatially correspondent with the viewer's real hand and is directly controlled by the instrumented glove device. With this capability, the

operator can pick-up and manipulate virtual objects that appear in the surrounding virtual environment. Similarly, in VIEW system developed data-management environments, multiple windows of information and simulated control panels are positioned, sized and activated by simply manipulating the virtual objects in 3-space. In coordination with the connected speech recognition technology, the hand and arm gesture information can also be used to effect and/or direct speech-indicated actions in the synthesized or remote environment (e.g., control of robotic arms and end-effectors, and associated control of remote camera viewpoints) (Fisher, 1986; Foley, 1987).

RESEARCH SCENARIOS

Telerobotics

Control of autonomous and semi-autonomous telerobotic devices and vehicles requires an interface configuration that allows variable modes of operator interaction ranging from high-level, supervisory control of multiple independent systems to highly interactive, kinaesthetic coupling between operator and remote system. An appropriate interface for supervisory control modes will provide the operator with multiple viewpoints of the remote task environment in a multi-modal display format that can be easily distributed and reconfigured according to changing task priorities. For remote operations that cannot be performed autonomously, the interface will need capability to quickly switch to interactive control. In this telepresence mode, the operator will require a sufficient quantity and quality of sensory feedback to approximate actual presence at the remote task site.

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

Dataspace

Similar to interface requirements for supervisory control of telerobotic systems, the operator interface for large-scale integrated information systems such as Space Station will also require a range of interaction modalities for operator intervention in the case of system degradation or conflicts in resource allocation. Efficient supervision of these automated systems will depend on highly graphic, multi-dimensional status representations of the numerous sub-systems in a format that can be easily monitored in parallel with other mission related tasks such as planning and scheduling, telerobot supervision, and communication activities. In the event of system conflict or malfunction, the interface environment should enable natural interaction with multi-modal, knowledge-based warning and advisory systems.

Advanced data display and management concepts for this task environment are being developed with the virtual environment display technology. Current investigations include use of the system to create a display environment in which data manipulation and system monitoring tasks are organized in virtual display space around the operator. Through speech and gesture interaction with the virtual display, the operator can rapidly call up or delete information windows and reposition them in 3-space. Three-dimensional sound cues and speech-synthesis technologies are used to enhance the operators overall situational awareness of the virtual data environment. The system also has the capability to display reconfigurable, virtual control panels that respond to glove-like tactile input devices worn by the operator. Major research issues include development of multi-modal data access and manipulation guidelines; concepts for multi-dimensional, graphic representation of knowledge-based information systems; and definition of interface configurations for shared workspace environments in collaborative systems management.

DISCUSSION

Research Directions

Projected near term technology developments for the VIEW system include improved and distortion-corrected computer generated imagery processed at significantly higher frame rates; High resolution, color display elements; Integration of 3D auditory display technology for spatially correspondent, synthesized sound cueing; Tactile feedback capability; Multiple viewpoint display

integration; And multiple VIEW display stations for shared workspace configurations in which each user is graphically present and interactive in the Virtual Environment.

Conclusions

Unlike most technology for 360-degree visual simulation environments, the virtual environment display system does not make use of large, expensive, special purpose projection configurations. The described system is portable and low-cost without large space and equipment requirements. In comparison to other research efforts in head-mounted displays, this system is unique in presenting a stereoscopic image that closely matches human binocular vision capabilities and in its configuration with state-of-the-art auditory, speech and tactile input technology.

The VIEW systems' capabilities for providing a highly graphical, uniform interface for varying task environments and level of interaction may reduce operator workload and training requirements and increase productivity, and its use of multi-modal input channels and augmentation of visual displays to provide unobtrusive redundancy may increase accuracy and efficiency in human-machine interaction. Also, multiple viewpoint presentation of task related information in a 360-degree, dynamic stereoscopic display appears to increase situational awareness and effectiveness in monitoring cognitively demanding spatial tasks such as construction and proximity operations or spatial representations of knowledge-based systems and activities.

REFERENCES

- Fisher, S.S. (1986) Telepresence Master Glove Controller for Dexterous Robotic End-Effectors, Advances in Intelligent Robotics Systems, D.P.Casasent, Editor, Proc. SPIE 726, 1986.
- Foley, J.D. (1987), Interfaces for Advanced Computing. Scientific American, 257(4), 126-135.
- Wenzel, E.M., Wightman, F.L., & Foster, S.H. (1988) A Virtual Acoustic Display for Conveying Three-Dimensional Information. Proc. Hum. Fac. Soc., 1988.