A Synthetic Environment System for Planning, Education and Data Visualization

John Sadowsky
Ian E. Feldberg
David M. Weintraub
Steven D. Diamond

The Johns Hopkins University Applied Physics Laboratory
Laurel, Maryland 20723-6099

Abstract

An inexpensive workstation that creates an immersive environment for several people to see and manipulate computer generated 3D graphic objects would have many innovative and important applications, such as planning and brainstorming (3D viewgraphs), education (3D chalkboard), and data visualization. We have been developing a PC-based system with a multimedia architecture. Each participant uses a head-mounted display with a video camera at each eye, and mixes the live video with computer generated graphics to create a heads-up display of the common scene with added virtual objects. We present the architecture, design problems and solutions, and application experiences.

1. Introduction

The technology of virtual reality (VR) has been promising significant innovations and revolutionary changes in the ways in which people interact with computers and computer-generated data environments for over twenty years. The ultimate goal of VR systems, whether or not it is completely achievable, is to provide devices, sensors, and computer processing to remove the artificial interfaces between the operator and the computer (e.g., keyboard, mouse, and monitor) and make it possible for direct contact and experience with the computer generated data space. Although research and development has produced several clever and useful VR systems, major advances have been difficult to achieve, in part because of the high expense of the technology and in part because of the basic science that must be investigated and understood to build compelling, immersive synthetic environments [1,2,3,4].

Over the past several months, we have been investigating architectures for providing an inexpensive VR experience, in situations in which the computer generated environment is to be part of the real world. In the field, this sort of application is often called heads-up display or environmental overlay. The objective is to allow the user to see computer generated objects or enhancements in the natural world, rather than to place the user in a totally artificial environment. In proposed heads-up systems, the user would wear a set of transparent goggles onto which could be projected computer generated images. As an example, such a system could support aircraft maintenance workers by overlaying computer stored and generated schematic diagrams of the underlying electronic circuitry or substructure systems over the hull of the aircraft being serviced.

Our approach is different in that the combining of computer generated and real world environments is performed via video mixing, rather than by projection onto a goggles surface. In this way, image registration and integration problems can be solved by digital image processing within the video mixing function. A suitable head-mounted display was developed at Johns Hopkins University and has been combined with appropriate devices and software for head position and orientation tracking, computer graphics and 3D display, and video mixing.

The architecture for the system will be presented in the following sections. Design decisions made to reach this particular architecture will be discussed and future enhancements will be proposed. The principal design goals were to build a usable VR system that is relatively inexpensive and of sufficient quality to investigate the feasibility and utility of such systems for a variety of applications.

Some of the applications that we are planning for the system are in the areas of military planning and wargaming, scientific education, and data visualization. The applications will be discussed in this paper.

2. System architecture

An important design goal of the system is that it be as inexpensive as can be achieved with existing devices, such that the performance requirements can be met. The performance requirements can be somewhat relaxed, as our purpose is to have a laboratory instrument to investigate feasibility and utility of such systems; however, the system must be of sufficient quality that the promise of the approach for the proposed applications is readily apparent.

In most current VR systems, the two most expensive components are the computer platform and the head-mounted display. For reasonably high fidelity computer graphics that can be generated in real time, high-end graphics workstations, such as mid-to-high range Silicon Graphics Workstations, are usually employed as the computer platform. As of mid-1993, when the architecture was being defined, a moderate resolution (of the order of 1000 pixels on a side) monochrome, head-mounted display with head tracking capability was available at a cost of between $60,000 and $80,000. In our applications, especially for education, these were significantly beyond budget.

High end graphics workstations, capable of generating scenes with 10,000 to 20,000 polygons per second, also offer more computational power than our applications require. Because the background scenery uses live video, the computer is only required to produce specific objects to be inserted into the environment. Thus, it is estimated that the system need only be able to generate 5,000 to 7,000 polygons per second. A fast PC with the appropriate graphics accelerator boards is capable of achieving this power. Thus, we have selected a high end PC, with a 90 MHz Pentium processor as the computing platform, and enhanced it with two SPEA Fire graphics accelerator boards, each board to provide a graphics input for a different eye in the stereo head-mounted display. In anticipation of texture mapping requirements, the boards were further enhanced with additional memory.

The Johns Hopkins University Medical School, in the Low Vision Center of the Wilmer Eye Clinic, has developed a head-mounted display, called the Low Vision Enhancement System (LVES) [5]. This device was invented to aid patients with severe chronic visual impairments. Simply stated, the LVES consists of two lightweight video cameras, mounted in front of the eyes, to detect the actual scene in front of the patient. Through an electronic and optical path, these video inputs are sent to two CRT displays that are reflected into the left and right eyes of the patient (see figure 1.) The video images are intercepted within the path and put through an image enhancement processor, to increase contrast and to detect and emphasize edges, so that low vision patients can perform basic visual tasks such as face recognition. Because of the demand for such a device, the cost is significantly lower than other head-mounted displays, currently at $5,000, but expected to decrease to between $3,000 and $4,000 in the near future. Although the displays in the current version are monochrome, the CRT technology employed provides a high resolution performance which is comparable to the highest quality currently available from digital devices.

The LVES is ideally suited for the application to the synthetic environment system. In our system, the image processor of the LVES is replaced with a video mixer that combines the input from its front-end cameras with the computer generated objects to produce a synthetic environment.

For the software to drive the system, we have selected a commercial package, WorldToolKit, from the Sense8 Corporation. This software is in the form of a collection of C-library routines, with some demonstration drivers provided, to generate polygonal representations of graphics objects, determine stereoscopic projections of the objects in left- and right-eye points-of-view, and incorporate position information from a variety of different tracking devices. In addition, WorldToolKit has the appropriate hooks to employ the graphics capabilities of the SPEA Fire boards. Although the library can be used in C language applications, it is more natural to develop object-oriented programs to create and manipulate the graphics objects. The library is easily used by C++ drivers as well.

Position and orientation tracking is accomplished using the Fastrak device, sold by the Polhemus Corporation. This device consists of a transmitter, which sequences through the generation of three orthogonal magnetic fields AC, and one or more receivers. The receivers each consist of three orthogonal coils of wire. In the magnetic field, an electric field is induced in the coils and the strength of this field in each coil, with respect to which magnetic field is being generated, provides six positional measurements for the receiver: the x, y, and z location of the receiver with respect to the transmitter, and the roll, pitch, and yaw angles of the receiver with respect to a calibrated set of axes.

In our architecture, the tracking device uses two receivers, one mounted to the LVES and the other mounted to a mouse, to act as a three-dimensional mouse. The particular tracking device was selected because of its low cost and the ease with which the Fastrak device can be driven on a PC platform. The primary limitation with this device is range; the device loses functionality at a range of between 5 and 10 feet between transmitter and receiver. For some of our applications, this could require the selection of another tracking device for operational use.

The present system uses an external video mixer, although a future goal is to perform the mixing with sophisticated image processing algorithms on an internal processor in the computer. This design decision was made largely in the interest of time, as it was desired to have an immediate demonstration of the LVES and video mixed environmental overlay. Thus, inexpensive commercial video mixers (one per eye) were purchased for the project.

The method of mixing used is chroma-key. The computer generated objects are built on a solid color, neutral background, using a color which would not be used in the generated objects. At present, this is not a restriction as the displays are monochrome and color is not an important consideration. In the video mixer, all pixels of the selected background color are converted to the corresponding pixels from the digitized live video image from the LVES front-end camera.

Appropriate signal converters (e.g., from the computer-generated RGB to the standard video NTSC format for input to the mixer) are inserted at the appropriate locations to complete the architecture. A schematic representation of the architecture is included as figure 2. The total cost of the entire system is approximately $30,000, however, it is anticipated that production versions of the system would be less than $20,000 per unit.

For many applications of such a workstation, it is necessary for the workstations to be able to share data and situations. For example, in a planning meeting of several people interacting with each other and with computer generated graphics, participants must be able to see each other and see the virtual objects as well. The shared experience of a meeting is accomplished via the live video from the head-mounted cameras---participants will see each other in the graphics background without the need for the workstations to acquire and share this information. The computer generated objects, however, must reside in the same locations and orientations in the synthetic environments of each participant, at the particular perspective and viewpoint of the participant. Because these objects are virtual, existing only within the domain of the computer, information about the object and modifications to the object must be shared among the workstations.

A simple solution to this problem would be to create and run the same application software on all of the workstations, but with separate position trackers feeding position 6-tuples independently to each workstation. As long as objects are not modified in a particular session, this simple approach produces the illusion that each participant is seeing the same virtual object from the appropriate viewpoint. For example, suppose one is in a mathematics classroom in which the instructor wishes to show a complex 3D topological space, and students are permitted to walk about the classroom to see the surface from different perspectives. If one desires to modify the object, however, then the modifications must be communicated to all of the workstations used by the instructor and the students.

Modifications to objects are made with the use of the three-dimensional mouse, built from a PC mouse by the addition of a position tracking receiver. (It would not suffice to place a receiver from each workstation on the same mouse, because the mouse button clicks must be communicated as well to each computer.) The workstation communications are accomplished by a serial link, which is straightforward to program. It is planned to provide Ethernet connections between PC's and have changes to the objects communicated as messages to all of the other stations. In this way, more complex processing and modifications to virtual objects, using the keyboard or other input devices, can be performed at any particular workstation and perceived by all participants.

3. Applications

The objectives of our project are to develop experience and expertise with the tools of VR, to build particular problem applications with general purpose synthetic environment systems, and to assess the feasibility and effectiveness of these problem applications. Thus, for example, if one application is to produce virtual three-dimensional models to be used in a mathematics classroom, then we want to determine not only that we can simulate geometric constructions and spaces, but that the use of such virtual models has a significant advantage in student performance and curriculum flexibility, compared with more traditional use of perspective pictures, solid models, and creating mental images. The cost of a synthetic environment system must be justified by the level of performance improvement, compared with more traditional methods for visualization.

To illustrate the effectiveness of the system, we have identified three applications. They are: military planning and situational awareness, three-dimensional chalkboard for mathematics education, and workstation for three-dimensional data visualization.

An important VR system of great use and benefit for military training, strategic and tactical planning, and for effectiveness assessment is the Distributed Interactive Simulation (DIS). Evolving from the SIMNET system developed by the Defense Advanced Research Projects Agency, DIS refers to a set of standards and protocols for data provided to and by computer simulators that communicate over a distributed network. As an example, if it is desired to have a tank simulator in California, a helicopter simulator in Maryland, and a second helicopter simulator in Denver all participate in a wargaming exercise for a training purpose, then one would want to have all three simulators follow the same protocols and formats for sending and receiving data about position, motion, shooting, etc. to a common network. In this way, the tank can see the helicopters when they are in view, the helicopters can likewise see each other and the tank, accurate targeting and shooting can be experienced by the appropriate simulators, and everyone can participate in a common exercise in a synthetic battlefield that virtually exists in a distributed network of computers.

An obvious extension of the DIS/SIMNET concept would be to allow military commanders, training instructors, or students of battlefield assessment to be able to see the overall state of the wargame by intercepting and processing all of the data being passed across the network. The problem is how best to present and analyze this data to have a realistic and useful representation of the wargame. One way would be to display on a computer monitor a perspective image of the state of the battlefield, representing the participating weapons systems simulators by appropriate icons or models. All observers would have a quick and intuitive representation of the situation. The picture on the monitor could be updated in real time to see a battle as it evolves. Modifications to the image on the screen, for example to change the observers' viewpoint, to remove obstacles for better pictures, or to add other data and pertinent information, could easily be achieved with keyboard entries.

As the battlefield is three-dimensional, an alternate method for displaying the DIS data would be to use the VR system described in this paper. Here, the observers, training instructors, and commanders would be in the same room, a Wargaming Analysis Laboratory, for example, wearing the head-mounted displays. Interaction and discussion are facilitated because everyone sees each other as in any meeting or conference. In addition, the battle itself is unfolding on a virtual battlefield in the center of the room, providing all participants with an aerial view of the situation. In addition, relevant data and other visualizations can be superimposed on the displays of all participants.

A second area of application being developed for this system is education, especially mathematics education. The instructor and students would wear the head-mounted displays and the system would act as a three-dimensional chalkboard onto which can be drawn virtual images of multidimensional calculus surfaces, topological spaces, combinatorial graphs, and other geometric visualizations of mathematical objects. As a teaching device, the VR system satisfies a need expressed by instructors. As a research instrument, this application will be used to understand the potential advantage of three-dimensional visualization of abstract concepts for developing mathematical intuition.

In addition, we are developing the software to generate and present three-dimensional data, such as time-frequency surfaces from signals, in a setting for viewing in different perspectives and for collaborative discussions. This application for data visualization also has promise for a variety of important medical image understanding, such as 3D MRI image reconstruction and 3D mammography.

4. Conclusions

Once the components and devices for the system described in this paper were gathered together, it became clear that producing a compelling illusion of VR is quite difficult and would require an enormous amount of research, in the scientific fields of human sensory processing, in the cognitive fields of human performance and human factors, and in the obvious disciplines of computer science. Putting together a system with which to experiment is certainly a necessary first step, but much work has to be done. The current project is in the process of building the particular applications described, especially for the military wargaming system and for the three-dimensional chalkboard.

In addition, we are anticipating near-term enhancements to the devices used in the system. The supplier of the head-mounted display is planning to have a color display version by mid-1995 and a version that employs eye tracking (for the application to low vision as well) by the end of 1996. Although the heads-up display approach being used does not require realistic sound generation and auditory processing, we are planning to enhance the manipulative abilities of the system by incorporating data gloves and force feedback for a sense of touch and haptic sensation. An application of the system to biological wet laboratory (e.g., dissecting the virtual frog) has motivated this avenue of research.

For a more realistic mixing of the live video and computer generated graphics, we are investigating image understanding and image registration algorithms to apply to the video mixer. As an example, the current system paints the video image over the field of the display. One cannot see a hand touch and grab an object. With a dataglove to provide coarse information about the position of the hand and object recognition capability, to find and track the hand in the live video, the computer generated object can be modified so that parts of it appear to be occluded by the hand. Efficient and accurate algorithms for such image processing are an area for research.

5. References

[1] Aukstakalnis, S. and Blatner, D., *Silicon Mirage*, Peachpit Press, Berkeley, California, 1992

[2] Kalawsky, R., *The Science of Virtual Reality and Virtual Environments*, Addison-Wesley, Reading, Massachusetts, 1993

[3] Pimentel, K., *Virtual Reality: Through the New Looking Glass*, Intel/Windcrest, New York, 1993

[4] Sadowsky, J. and Massof, R., "Sensory Engineering: the Science of Synthetic Environments," in *Johns Hopkins University Technical Digest*, Vol. 15, No. 2, June 1994, special issue on Synthetic Environments, pp. 99-109

[5] Massof, R., Rickman, D., and Lalle, P., "Low Vision Enhancement System," in *Johns Hopkins University Technical Digest*, Vol. 15, No. 2, June 1994, special issue on Synthetic Environments, pp. 120-125

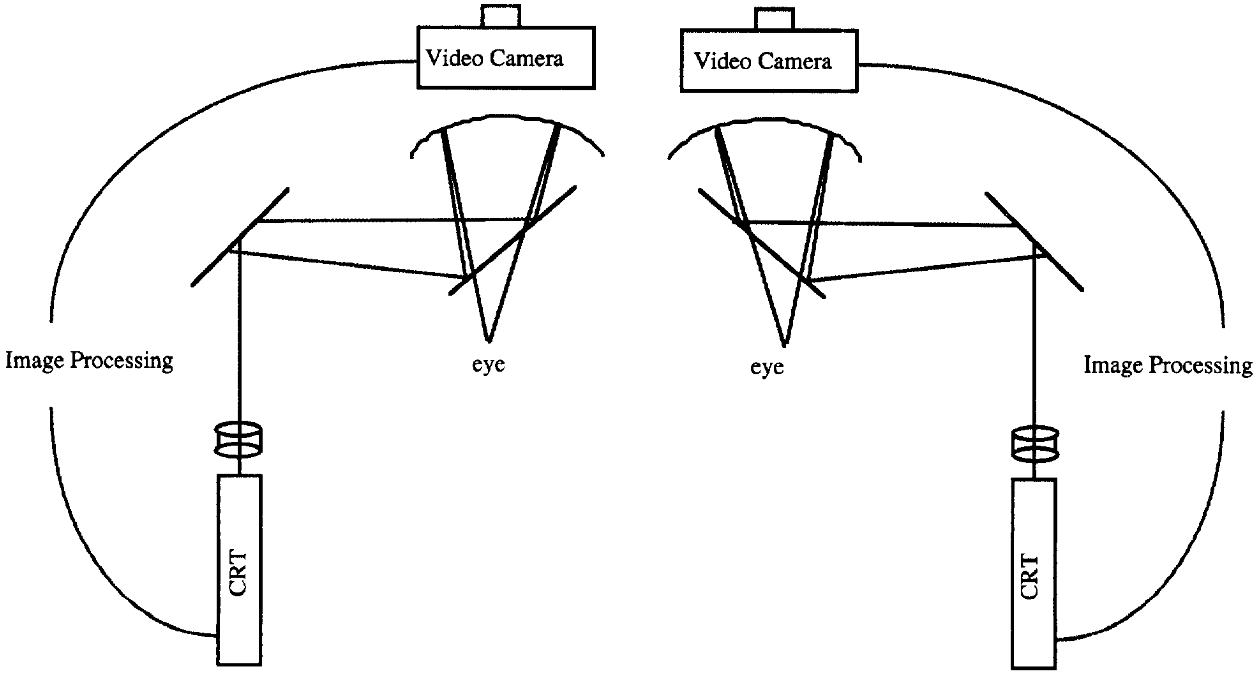

Figure 1. Schematic Diagram of the LVES Head-mounted Display subsystem used for the synthetic environments workstation. The image processing is replaced with a video mixer with additional input from the computer graphics generation subsystem

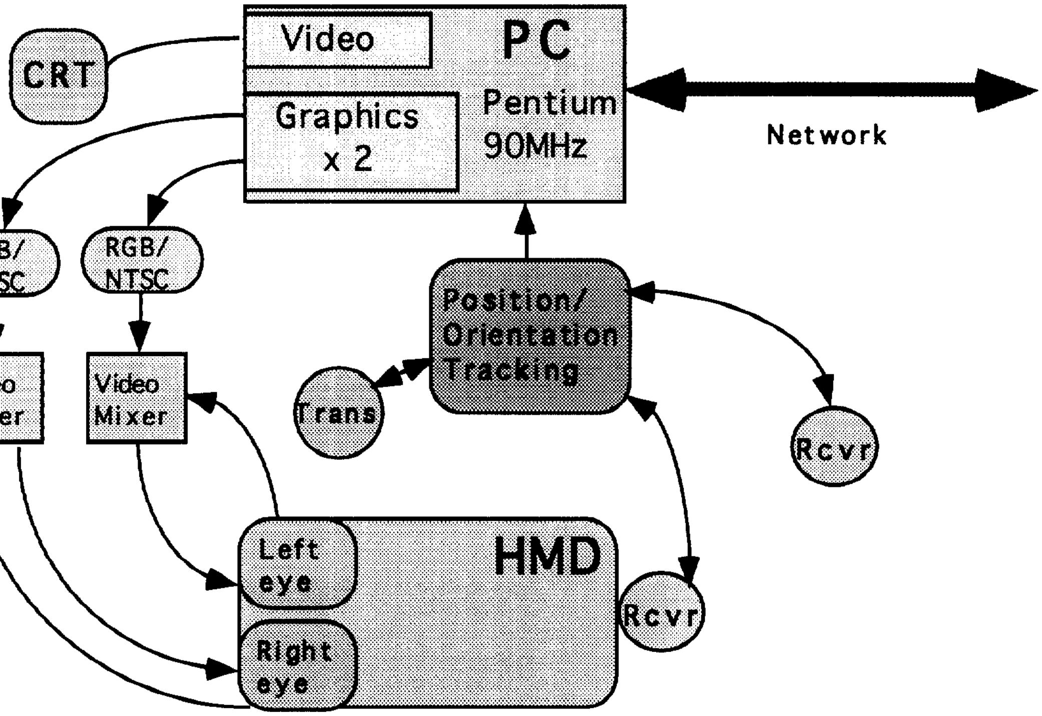

Figure 2. Synthetic Environments Workstation Architecture