

Surround-Screen Projection-Based Virtual Reality: The Design and Implementation of the CAVE

Carolina Cruz-Neira[†]

Daniel J. Sandin

Thomas A. DeFanti

Electronic Visualization Laboratory (EVL)

The University of Illinois at Chicago

Abstract

This paper describes the CAVE (CAVE Automatic Virtual Environment) virtual reality/scientific visualization system in detail and demonstrates that projection technology applied to virtual-reality goals achieves a system that matches the quality of workstation screens in terms of resolution, color, and flicker-free stereo. In addition, this format helps reduce the effect of common tracking and system latency errors. The off-axis perspective projection techniques we use are shown to be simple and straightforward. Our techniques for doing multi-screen stereo vision are enumerated, and design barriers, past and current, are described. Advantages and disadvantages of the projection paradigm are discussed, with an analysis of the effect of tracking noise and delay on the user. Successive refinement, a necessary tool for scientific visualization, is developed in the virtual reality context. The use of the CAVE as a one-to-many presentation device at SIGGRAPH '92 and Supercomputing '92 for computational science data is also mentioned.

Keywords: Virtual Reality, Stereoscopic Display, Head-Tracking, Projection Paradigms, Real-Time Manipulation

CR Categories and Subject Descriptors: I.3.7 [Three-Dimensional Graphics and Realism]: Virtual Reality; I.3.1 [Hardware Architecture]: Three-Dimensional Displays.

1. Introduction

1.1. Virtual Reality Overview

Howard Rheingold [11] defines virtual reality (VR) as an experience in which a person is “surrounded by a three-dimensional computer-generated representation, and is able to move around in the virtual world and see it from different angles, to reach into it, grab it, and reshape it.” The authors of this paper prefer a definition more confined to the visual domain: a VR system is one which provides real-time viewer-centered head-tracking perspective with a large angle of view, interactive control, and binocular display. A competing term, *virtual environments* (VE), chosen for “truth in advertising” [1], has a somewhat grander definition which also correctly encompasses touch, smell, and sound. Although VE is part of the CAVE acronym, we will use the initials VR herein to conform to mainstream usage.

Several common systems satisfy some but not all of the VR definition above. Flight simulators provide vehicle tracking, not head tracking, and do not generally operate in binocular stereo. Omnimax theaters give a large angle of view [8], occasionally in stereo, but are not interactive. Head-tracked monitors [4][6] provide all but a large angle of view. Head-mounted displays (HMD) [7][13] and BOOMs [9] use motion of the actual display screens to achieve VR by our definition. Correct projection of the imagery on large screens can also create a VR experience, this being the subject of this paper.

Previous work in the VR area dates back to Sutherland [12], who in 1965 wrote about the “Ultimate Display.” Later in the decade at the University of Utah, Jim Clark developed a system that allowed wireframe graphics VR to be seen through a head-mounted, BOOM-type display for his dissertation. The common VR devices today are the HMD and the BOOM. Lipscomb [4] showed a monitor-based system in the IBM booth at SIGGRAPH '91 and Deering [6] demonstrated the Virtual Portal, a closet-sized three-wall projection-based system, in the Sun Microsystems' booth at SIGGRAPH '92. The CAVE, our projection-based VR display [3], also premiered at SIGGRAPH '92. The Virtual Portal and CAVE have similar intent, but different implementation schemes.

To distinguish VR from previous developments in computer graphics, we list the depth cues one gets in the real world.

- 1 Occlusion (hidden surface)
- 2 Perspective projection
- 3 Binocular disparity (stereo glasses)
- 4 Motion Parallax (head motion)
- 5 Convergence (amount eyes rotate toward center of interest, basically your optical range finder)
- 6 Accommodation (eye focus, like a single-lens reflex as range finder)
- 7 Atmospheric (fog)
- 8 Lighting and Shadows

Conventional workstation graphics gives us 1, 2, 7, and 8. VR adds 3, 4, and 5. No graphics system implements accommodation clues; this is a source of confusion until a user learns to ignore the fact that everything is in focus, even things very close to the eyelash cutoff plane that should be blurry.

The name of our virtual reality theater, “CAVE,” is both a recursive acronym (CAVE Automatic Virtual Environment) and a reference to “The Simile of the Cave” found in Plato’s *Republic* [10], in which the philosopher discusses inferring reality (ideal forms) from projections (shadows) on the cave wall. The current CAVE was designed in early 1991, and it was implemented and demonstrated to visitors in late 1991. This paper discusses details of the CAVE design and implementation.

[†] 851 S. Morgan, Room 1120 SEO (M/C 154). Chicago, IL 60607-7053.
E-mail: cruz@bert.eecs.uic.edu

Permission to copy without fee all or part of this material is granted provided that the copies are not made or distributed for direct commercial advantage, the ACM copyright notice and the title of the publication and its date appear, and notice is given that copying is by permission of the Association for Computing Machinery. To copy otherwise, or to republish, requires a fee and/or specific permission.
©1993 ACM 0-89791-601-8/93/008...\$1.50

1.2. CAVE Motivation

Rather than having evolved from video games or flight simulation, the CAVE has its motivation rooted in scientific visualization and the SIGGRAPH '92 Showcase effort. The CAVE was designed to be a useful tool for scientific visualization. Showcase was an experiment; the Showcase chair, James E. George, and the Showcase committee advocated an environment for computational scientists to interactively present their research at a major professional conference in a one-to-many format on high-end workstations attached to large projection screens. The CAVE was developed as a "Virtual Reality Theater" with scientific content and projection that met the criteria of Showcase. The Showcase jury selected participants based on the scientific content of their research and the suitability of the content to projected presentations.

Attracting leading-edge computational scientists to use VR was not simple. The VR had to help them achieve scientific discoveries faster, without compromising the color, resolution, and flicker-free qualities they have come to expect using workstations. Scientists have been doing single-screen stereo graphics for more than 25 years; any VR system had to successfully compete. Most important, the VR display had to couple to remote data sources, supercomputers, and scientific instruments in a functional way. In total, the VR system had to offer a significant advantage to offset its packaging. The CAVE, which basically met all these criteria, therefore had success attracting serious collaborators in the high-performance computing and communications (HPCC) community.

To retain computational scientists as users, we have tried to match the VR display to the researchers' needs. Minimizing attachments and encumbrances have been goals, as has diminishing the effect of errors in the tracking and updating of data. Our overall motivation is to create a VR display that is good enough to get scientists to get up from their chairs, out of their offices, over to another building, perhaps even to travel to another institution.

1.3. CAVE Design

The CAVE we exhibit at conferences is a theater 10'x10'x10' made up of three rear-projection screens for walls and a down-projection screen for the floor, as shown in Figure 1. (Our development system at EVL is actually 7'x7'x7' due to ceiling height limitations.) Projectors throw full-color workstation fields (1280x512 stereo) at 120Hz onto the screens, giving between 2,000 and 4,000 linear pixel resolution to the surrounding composite image. Computer-controlled audio provides a sonification capability to multiple speakers. A user's head and hand are tracked with Polhemus or Ascension tethered electromagnetic sensors. Stereographics' LCD stereo shutter glasses are used to separate the alternate fields going to the eyes. Four Silicon Graphics high-end workstations create the imagery (one for each screen); they are tied to a fifth for serial communications to input devices and synchronization via fiber-optic reflective memory by Systran Corporation. The CAVE's theater area sits in a 30'x20'x13' room, provided that the projectors' optics are folded by mirrors. Conference use thus far has necessitated the building of a light-tight structure of this size on site to house the screens and projectors.

Goals that inspired the CAVE engineering effort include:

- 1 The desire for higher-resolution color images and good surround vision without geometric distortion.
- 2 Less sensitivity to head-rotation induced errors

- 3 The ability to mix VR imagery with real devices (like one's hand, for instance)
- 4 The need to guide and teach others in a reasonable way in artificial worlds
- 5 The desire to couple to networked supercomputers and data sources for successive refinement

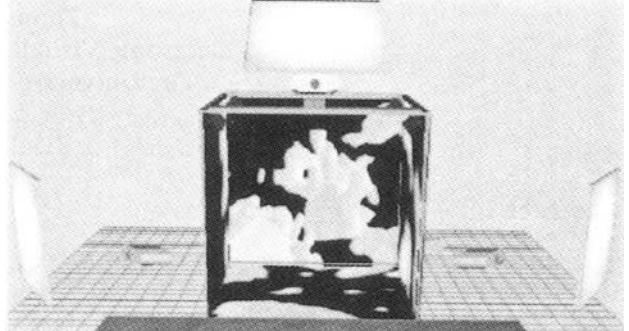


Figure 1: CAVE diagram. Graphics by Milana Huang, University of Illinois at Chicago

Significant barriers, now hurdled, include eliminating the lag inherent in common green video projector tubes, corner detailing, and frame accurate synchronization of the workstations; our solutions to these problems are described in detail in section 3. The electromagnetic trackers required building the CAVE screen support structure out of non-magnetic stainless steel (which is also relatively non-conductive), but nonlinearities are still a problem, partially because conductive metal exists on the mirrors and in the floor under the concrete. Wheelchairs, especially electric ones, increase tracker noise and non-linearities as well.

Unsolved problems to date include removing the tracking tether so the user is less encumbered, moving the shutters from the eyes to the projectors so cheap cardboard polarizing glasses can be used, incorporating accurate directional sound with speakers, and bringing down the cost. These, and other problems we've encountered, are described in section 6.

The implementation details fall mainly into two categories: projection and stereo. These will be presented next.

2. Projection Details

2.1. Cube Sides As Projection Planes

One rarely noted fact in computer graphics is that the projection plane can be anywhere; it does not have to be perpendicular to the viewer (as typical on workstations, the HMD, and the BOOM). An example of an unusual projection plane is the hemisphere (like in Omnimax theaters or some flight simulators). However, projection on a sphere is outside the real-time capability of the ordinary high-end workstation. And, real-time capability is a necessity in VR.

The CAVE uses a cube as an approximation of a sphere. This simplification greatly aids people trying to stand in the space, and fits the capabilities of off-the-shelf graphics and high-resolution projection equipment, both of which are made to create and project imagery focused on flat rectangles. The defects one encounters in attempting to build a perfect cube are fortunately within the range of adjustment by standard video projectors; in particular, keystone and pincushion corrections can be utilized. Thus, the ability to match projected images at the seams and corners is effectively perfect, with tuning effort.

2.2. Window Projection Paradigm

The most common computer graphics projection paradigm is the camera view. This type of projection simulates the way an image is captured on film, and includes the direction the camera is pointed and the focal length, position, and twist angle of the lens. In the camera paradigm, stereo is typically achieved by using two cameras; this is the technique used by the HMD and BOOM. The CAVE instead uses a window projection paradigm in which the projection plane and projection point relative to the plane are specified, thus creating an off-axis perspective projection.

Fortunately, the Silicon Graphics' Graphics Library (GL) [14] provides a window projection function. Since this function can also be performed by two shears and a standard perspective projection, or, alternatively, by a translation, a standard perspective projection and a translation back, the window projection function can easily be constructed from more primitive functions, if not available in another graphics library.

In the CAVE, the projection plane locations correspond to the locations of the actual walls. Therefore, as the viewer moves around in the environment, the off-axis stereo projection is calculated according to his/her position with respect to the walls (see Figure 2).

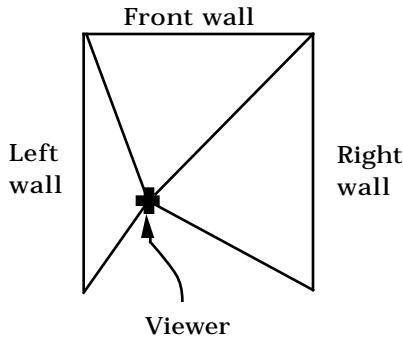


Figure 2: Off-axis projection

For the simplicity of the calculations, we assume that all the walls share the same reference coordinate system as shown in Figure 3. The origin of the coordinate system is placed in the center of the CAVE and it is a right-handed system with respect to the front wall. All the measurements from the trackers (position and orientation) are transformed to match this convention.

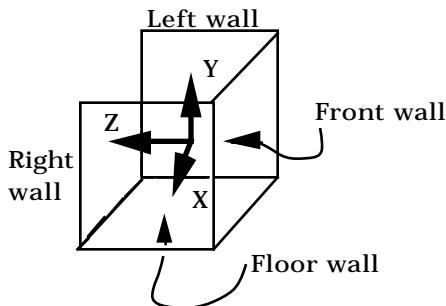


Figure 3: CAVE reference system.

Figure 4 shows a top diagram of the CAVE. The point Q' is the projection of the point Q . PP is the distance from the center of the CAVE to the front wall (5' for the 10'x10'x10' CAVE).

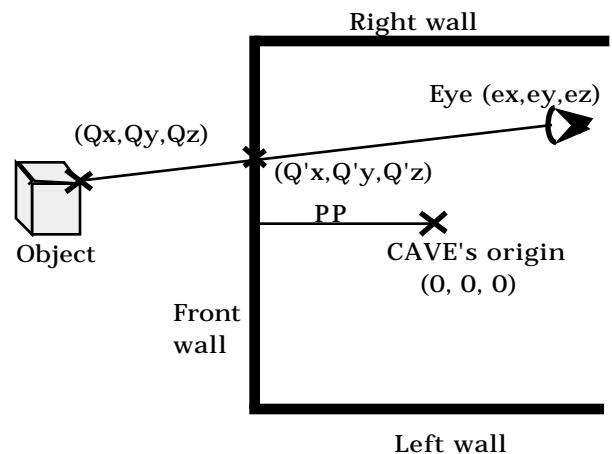


Figure 4: CAVE projection diagram

Using straightforward algebra and following the conventions in Figure 4, the projection Q' of a point $Q(Q_x, Q_y, Q_z)$ on the front wall is given by:

$$Q'_x = Q_x + \frac{(PP - Q_z)(ex - Q_x)}{ez - Q_z}$$

$$Q'_y = Q_y + \frac{(PP - Q_z)(ey - Q_y)}{ez - Q_z}$$

Thus, the general projection matrix is:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \frac{ex}{ez - PP} & \frac{ey}{ez - PP} & 1 & \frac{1}{ez - PP} \\ \frac{exPP}{ez - PP} & \frac{eyPP}{ez - PP} & 0 & \frac{ez}{ez - PP} \end{pmatrix}$$

One important issue to mention is that, in the CAVE, the eyes are not assumed to be horizontal and in a plane that is perpendicular to the projection plane. A clear example of this is a situation in which the viewer is looking at one of the corners of the CAVE with his/her head tilted. Our tracker is mounted on top of the stereo glasses; it is raised 5.5" from the glasses to minimize interference and centered between the eyes. From the values obtained from the tracker, and assuming an interpupillary distance of 2.75", we can determine the position of each eye and its orientation with respect to each one of the walls before applying the projection matrix.

The reader can easily derive the matrices for the other walls of the CAVE. Notice that, since the walls of the CAVE are at exactly 90° from each other, the viewer's position with respect to the other walls are:

Left wall: (e_z, e_y, e_x)
Floor wall: $(e_x, e_z, -e_y)$

Right wall: $(-e_z, e_y, e_x)$

3. Stereo Vision Details

3.1. Convergence

To achieve stereo vision in the CAVE, we, in principle, do two off-axis stereo projections per screen, one for each eye. We need to obtain information from the tracker to accurately place each eye. We assume that the center of rotation of the eye is close

enough to the nodal point (projection point) of the eye to not introduce significant error. Thus, as with other VR systems, where the eyes are looking does not enter into the calculations.

3.2. Frame Sequential Stereo

To get a different image to each eye, we use frame sequential stereo with synchronized shutter glasses. Infrared transmitters cause the lens for each eye to stay transparent for the proper 512 lines of the 1280x1024 image per screen, switching during vertical retrace time. We produce 120 fields per second, thus updating the whole image at 60Hz, producing a flicker-free image.

Note, however, that the green phosphor used in commercially available projection tubes has a persistence that is too long, so a user always sees both images anyway, destroying the stereo effect. Until Stereographics provided us with P43 coated green tubes by special order, we did our experiments (in 1991) in blue and red and shades of magenta. With luck, tube manufacturers will be motivated to add such tubes to their catalogs soon.

3.3. Distortion Correction

The HMD, BOOM, and monitor VR systems have significant geometric distortion inherent in their optics. Modern data projectors have extensive electronic adjustments to accurately correct geometric distortions.

3.4. Minimizing User Shadows

The three wall screens are rear projected so that the participants in the CAVE do not cast shadows. The floor is down projected so shadows are cast. We off-axis project the image from the front top instead of directly overhead, so the shadow of the user falls mainly behind him/her.

3.5. Frame Accurate Synchronization

Another problem we had to solve was the perfect synchronization of the screen updates. If the images are even one frame out of sync, the images in the corners crease and start to look sucked in like sofa cushions. We were unable to get adequate response from the UNIX system to synchronize within the 8ms needed, so (at the suggestion of Silicon Graphics staff) we went to reflective memory, a sort of shared cache arrangement among all the workstations. Reflective memory allows C-pointers to directly access chunks of memory, neatly bypassing the operating system. We intend to use the reflective memory for more sophisticated data sharing, including broadcasting of meshes, textures, and polygon lists. For now, however, reflective memory solves a nasty problem.

3.6. Edge Matching

Particular attention is paid to the edges and corners of the screen to avoid occlusion of stereo objects inside the room. We minimize the seams by stretching a 10'x30' plastic screen over 1/8" stainless steel cable under tension. This gives a seam of about a pixel or so in width, which can be seen but can also be easily ignored. Hence, the illusion of stereo in the CAVE is extremely powerful to the viewer. The floor butts up against the screen fairly perfectly (1/16") and presents no problem.

In the case of 3D movies and workstation screens, stereo objects in front of the screen (often the most interesting ones) have to stay pretty much centered. When a stereo object in front of a screen hits the edge (called "frame violation" in the jargon), it collapses the depth illusion since occlusion is a stronger depth cue than binocular disparity. The CAVE's screen edges are

basically out of view (one can see the tops of the screens, but they are high up) so the stereo objects can be anywhere.

We were amazed at how much the floor adds to the experience; a user can walk around convincing objects that are being projected into the room. Since the tracker provides six degrees of information, the user's head can tilt as well, a natural way to look at objects. The HMD provides this capability, but BOOM hardware does not.

3.7 Minimizing Occlusion by Participants

A user's hand can cause stereo violation if an object is between the eyes and the hand, a rare enough situation. People are very eager to resolve stereo violation whenever it's easy so, in these instances, the user simply moves his/her hand out of the way.

A much more serious situation occurs with multiple people in the CAVE. If someone gets in the way of another viewer and an object is supposed to be projected between the two of them, the stereo collapses. We avoid this by having a "teacher" or "guide" control the navigation, but let the "student" or "tourist" be tracked and stand in front, thereby getting the best stereo experience without first having to learn to be an expert navigator of the data space, whatever it is. At conferences, we often jam a dozen people at a time in the CAVE and try to keep the images in front of the crowd. Since people more or less have to stay still or move together, the VR experience for all, however limited, is nevertheless pleasing.

3.8. Motion Sickness

Seeing one's own body or those of other people may in fact be a good idea. Of 9,000 or so people who have been in the CAVE, two have experienced enough nausea to complain about it, a very low ratio (apparently) for VR [1]. We don't yet know why the CAVE doesn't make people nauseous; perhaps it is content related. Our images primarily have to do with scientific data that changes over time, not roller coaster type motions with fast tilting horizons typical of many VR applications. Another explanation may be our better coping with fast head rotation (see next section).

4. Quantitative Analysis of the Effect of Tracking Noise and Latency

4.1. Introduction

Different VR modes have different responses to errors in tracking viewer position. One reason for the differences depends on whether the projection plane moves with the viewer (as with BOOMs and HMDs) or not (in the case of the monitor and CAVE). A second reason is the difference in the distance of the projection plane to the eye, which distinguishes the monitor implementation from the CAVE's.

4.2. Rotation errors

Tracking errors can be resolved into displacement errors and rotation errors. Actual problems are often a combination of the two. In the monitor and CAVE paradigms, since the projection plane does not move with the viewer's position and angle, a rotation about the projection point in the eye creates zero error. In the HMD/BOOM paradigm, a given rotational tracking error produces the same magnitude of rotational error in the image, but of opposite sign. This is a serious problem if the user's head rotates quickly because the whole visual scene first rotates with the head and then steps back into the proper place.

4.3. Analysis of displacement errors in the CAVE and monitor paradigms

The effect of displacement error for both the CAVE and the monitor paradigms is illustrated in Figure 8. The displacement error in eye tracking is ΔP (in a plane parallel to the projection plane), the distance from the eye to the projection plane is PD , and the distance to the object is Z . $DISP$ is the distance error on the projection plane. α is the angular error.

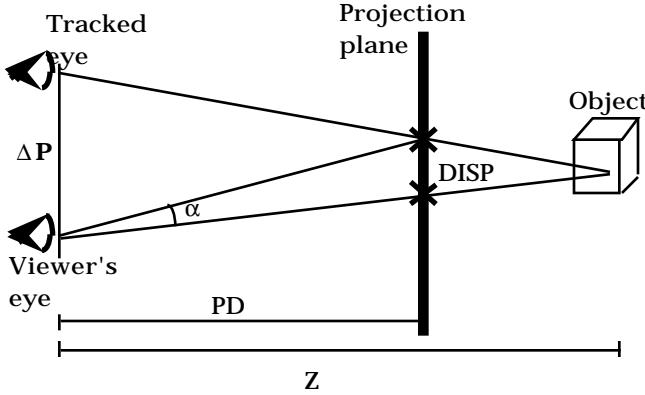


Figure 8: Effect of displacement error for both the CAVE and the monitor paradigms

$$DISP = \Delta P \left(\frac{Z - PD}{Z} \right)$$

$$\alpha = \arctan \left(\frac{DISP}{PD} \right)$$

$$\alpha \approx \frac{DISP}{PD} \text{ for small angles}$$

therefore,

$$(1) \alpha \approx \frac{\Delta P \frac{(Z - PD)}{Z}}{PD}$$

$$\text{For large } Z, \frac{(Z - PD)}{Z} \approx 1$$

therefore,

$$(2) \alpha \approx \frac{\Delta P}{PD}$$

$$\text{For small } Z, \frac{(Z - PD)}{Z} \approx -\frac{PD}{Z}$$

therefore,

$$(3) \alpha \approx -\frac{\Delta P}{Z}$$

For $Z = PD$ (when the object is on the projection plane),

$$\frac{(Z - PD)}{Z} = 0$$

therefore,

$$(4) \alpha = 0$$

Equation (1) represents the approximate angular error α for a displacement tracking error ΔP in the monitor and CAVE paradigms.

Equation (2) shows that the larger projection distance PD associated with the CAVE, as compared to the monitor, makes angular error α due to displacement ΔP smaller for large distances Z to the object viewed.

Equation (3) shows that for very small Z values, the monitor and CAVE have similar responses.

Equation (4) shows that when objects are on the projection planes of the monitor or CAVE, the angular error α due to displacement is zero.

4.4. Analysis of displacement errors in the BOOM and HMD

A similar analysis for the BOOM and HMD is indicated in Figure 9.

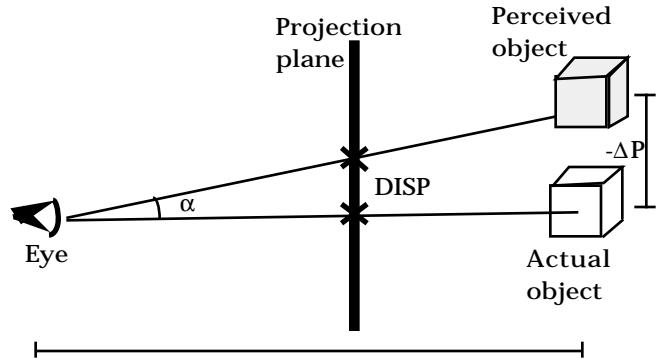


Figure 9: Effect of displacement error for both the HMD and the BOOM paradigms

A displacement error in tracking head position results in identical errors in both the eye position and the projection plane position. This results in a negative displacement of the object being viewed.

$$\alpha = \arctan \left(\frac{-\Delta P}{Z} \right)$$

For small angles,

$$(5) \alpha \approx \frac{-\Delta P}{Z}$$

Equation (5) shows that the angular error α is independent of the projection distance PD to the projection plane. Comparing equation (5) with (2), we see that the BOOM and HMD have less angular error α for displacement errors ΔP for large object distances Z than the CAVE/monitor models. Comparing equation (5) with (3), we see that the BOOM and HMD have similar angular errors α for small object distance Z .

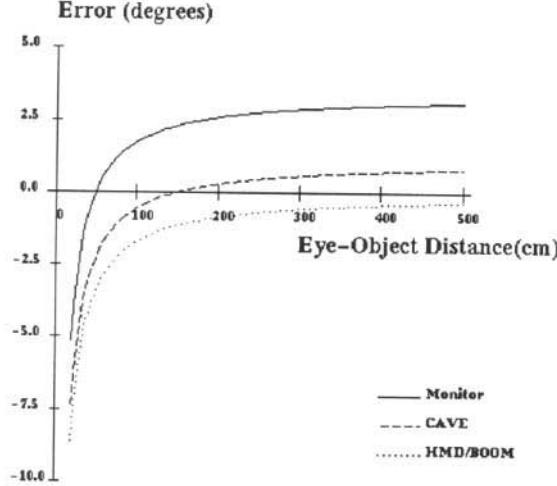


Figure 10: Angular error for a 3cm tracker displacement

Figure 10 graphs the angular error ΔP due to a tracker displacement error ΔP of 3cm for object distances Z . This case represents a tracking error due to latency of a person moving 30cm/second combined with a display rate of 10 frames/second. For large object viewing distances ($Z=500\text{cm}$), the HMD/BOOM have the best performance, the CAVE has 2-1/2 times the error, and the monitor has 9 times the error. For small object viewing distances ($Z=20\text{cm}$), the monitor has the best performance, and the CAVE and HMD/BOOM have only slightly worse error magnitudes.

4.5. Examples of combined rotation and displacement tracking errors

Normal head motions like nodding and panning involve both rotation and displacement of the eyes. The combined effect of these errors may be approximated by summing the individual angular errors a . The assumed projection distances PD for the monitor and 10' CAVE are 50cm and 150cm, respectively.

Figure 11 graphs the angular error a as a function of eye/object distance Z due to a head rotation (pan) of 90 degrees/second and a display rate of 10 frames/second. It is assumed that the eyes are 5cm from the center of rotation. For large Z , the CAVE is 43 times better than the HMD/BOOM and 4 times better than the monitor. For small Z , the CAVE and monitor are 6 times better than the HMD/BOOM.

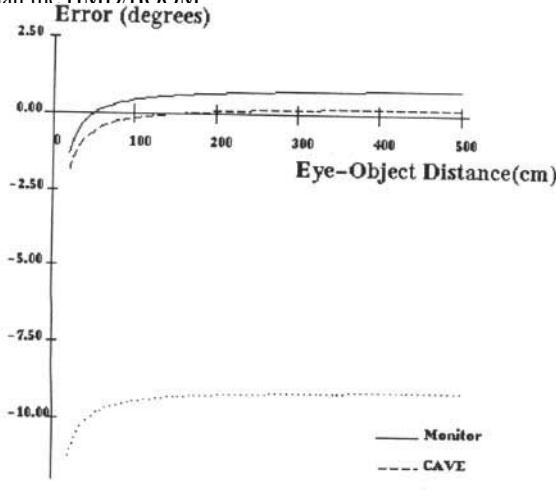


Figure 11: Tracking errors introduced by head panning

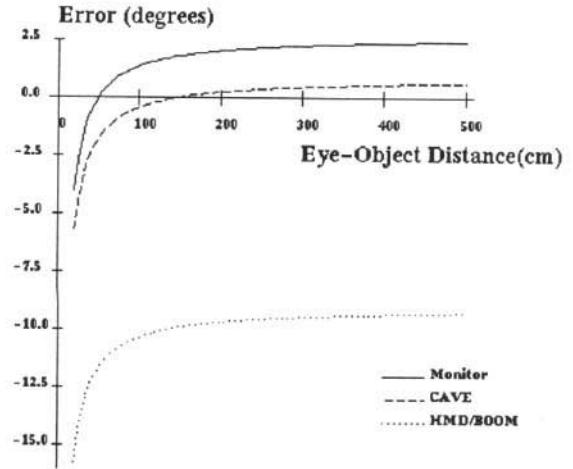


Figure 12: Tracking errors introduced by head nodding

Figure 12 graphs the angular error a as a function of eye/object distance Z due to a head rotation (nod) of 90 degrees/second and a display rate of 10 frames/second. It is assumed that the eyes are 15cm from the center of rotation. For large Z , the CAVE is 15 times better than the HMD/BOOM and 4 times better than the monitor. For small Z , the CAVE and monitor are 3 times better than the HMD/BOOM.

The examples above are all due to tracking errors caused by latency. Tracking errors from other sources, such as electrical interference, tend to be about an order of magnitude smaller, but the ratios are the same and we can draw the same conclusions. For the head-panning example in section 4.5, the problem was caused by normal head motion; if, however, we divide the angular error a by 20, we could interpret the graph as representing the case of a 0.5-degree tracking error combined with a tracking receiver mounted 5cm from the eye.

5. Successive Refinement

One benefit of the wrap-around screens in the CAVE is the potential for successive refinement of images. It is fair to say that we will never, in our lifetimes, have enough computing power to create complex models and display them in real time. Successive refinement trades off motion for time, freezing the image and filling it in, a now common computer graphics technique. Yet, one cannot freeze the image in a HMD without major disorientation. In the BOOM, successive refinement is possible but the user cannot look around. In the CAVE, one can navigate to a place in real time and then send off to a supercomputer for a highly detailed set of four images, still in stereo. When the images come back, the user can still pan around, although he/she cannot navigate while in this mode. The best stereo is achieved when looking in the last interactively tracked direction. Optimizing for this mode is the subject of active ongoing research.

Making VR usable in less-than-real-time situations is important. Supercomputers are essentially floating-point machines. One popular vector machine we use cannot create 1280x1024 pixel maps in real time because the floating-to-fixed conversions are done by non-vectorized subroutine calls (at three conversions, one for each pixel color component, it gets time consuming). There are no floating-point frame buffers for sale. In addition, the desire to transmit a 1280x1024 24-bit image to a workstation 60 times a second requires nearly 2 gigabits of network throughput! Multiply that by 4 for the CAVE screens. Since an

update rate of only 10 times a second is closer to VR industry standards, divide by 6, which results in a need for 1.25 gigabits/second. Clearly, we try to transmit polygon lists and meshes in floating point and let the workstation's graphics engine do its job whenever possible.

Naturally, it is important to consider more than image complexity; the basic science being computed often is extremely complex and will not respond in real time. Sometimes large stores of precomputed data are meaningful to explore; perhaps disk-based playback will be useful. The CAVE is a research resource now being used by scientists at the University of Illinois at Chicago, the National Center for Supercomputing Applications, Argonne National Laboratory, University of Chicago, California Institute of Technology, and the University of Minnesota. The overall goal is to match the capabilities of supercomputing, high-speed networking, and the CAVE for scientific visualization applications.

6. CAVE Shortcomings

6.1. Cost

The CAVE is big and expensive, although, given inflation, it is no more expensive than the PDP-11/Evans & Sutherland single-user display system was 20 years ago. Also, considering that up to 12 people can space-share the CAVE, the cost per person comes down in some circumstances. Cheap wall-sized LCD screens with low latency that one could stand on would be great to have, if they only existed. The desire for the rendering afforded by \$100,000 state-of-the-art graphics engines will not abate; however, current effects will be achievable at more modest cost as time goes on.

6.2. Ability to Project on All Six Sides of the CAVE

Six screens would make a better CAVE. We originally planned to do both floor and ceiling "rear" projections, which would have necessitated raising the CAVE structure 10'. A hole in the floor and a large sheet of strong glass or plastic would be a better solution, but not one easily achieved at conferences or universities.

A rear screen for the fourth wall might be possible, although the details for human entrance and exit would have to be worked out, especially if the cable-stretched screen technique were used. Four screens work very well, yielding large surround views for both panning actions and looking down. Consequently, objects inside the room can be walked around and virtually beg to be touched.

6.3. Light Spillage

One problem is the light spillage from the "screen" on the floor (the wall screens are fortunately not very reflective). Our floor screen is simply a painted floor board; the floor paint was quickly chosen by using the color-matching computer at the local paint distributor to duplicate the wall screens' color as a first approximation. The only time there would be a problem having one screen brighter than the others would be when the center of interest is not an object on the brightest screen, an unusual case. Very bright screens all around do tend to reduce image contrast somewhat, but this, too, has not been an issue. Naturally, good graphic design optimizes for the strengths and weaknesses of any medium.

6.4. Utilizing the CAVE Medium to Its Full Potential

The CAVE, like Omnimax, represents a different visual paradigm: inside out instead of outside in. From working with

students and colleagues, we realize that getting people to design visualizations and think in terms of inside-out is difficult, especially since the CAVE simulator used in the early stages of application development has an outside-in presentation on the workstation screen. Nonetheless, it is a concept into which it is fairly easy to incorporate data.

6.5. Fragility

The CAVE is not "museum hardy." The screens, tracker, and glasses are not kid-proof, thereby limiting use in museums, malls, arcades, and so on. More research is needed.

6.6. New Control Paradigms

As the computing community went from command-line terminals to 2D raster systems, the pull-down menu and mouse provided an alternative to the command line and keyboard. The CAVE has not produced any significant new control paradigms to date, although "step-on" menus have been proposed. One graduate student (Randy Hudson) has achieved a nice way to control rotation by having the user stroke a barely perceptible tessellated wireframe sphere with his/her hand. We look forward to the challenge of finding the next control models and encourage anyone with ideas to come and discuss collaboration.

6.7. Directional Sound

Another issue to address is the effective implementation of directional sound. In theory, with speakers in all corners, one should be able to achieve good directionality with the proper audio synthesis gear. In practice, however, sound localization is compromised by reflections off the screens.

6.8. Ability to Document

The CAVE is very hard to photograph. Imaginations soar when readers are presented with excellent suggestive 2D photos of other VR devices in use. We have not been able to compete in this domain. However, the CAVE and monitor are both amenable to video documentation if the tracking device is attached to the camera and the interocular distance is adjusted to zero.

7. Conclusions

The CAVE has proven to be an effective and convincing VR paradigm that widens the applicability and increases the quality of the virtual experience. The CAVE achieves the goals of producing a large angle of view, creating high-resolution (HDTV to twice HDTV) full-color images, allowing a multi-person (teacher/student or salesperson/client) presentation format, and permitting some usage of successive refinement. Furthermore, the flatness of the projection screens and the quality of geometric corrections available in projectors allow presentations of 3D stereo images with very low distortion as compared to monitor-based, HMD, and BOOM VR systems. The user is relatively unencumbered given that the required stereo glasses are lightweight and the wires to the head and hand trackers for the tracked individual are very thin. Since the projection plane does not rotate with the viewer, the CAVE has dramatically minimized error sensitivity due to rotational tracking noise and latency associated with head rotation, as compared to the HMD and BOOM.

At SIGGRAPH '92 and Supercomputing '92, more than a dozen scientists, in fields as diverse as neuroscience, astrophysics, superconductivity, molecular dynamics, computational fluid dynamics, fractals, and medical imaging, showed the potential of the CAVE for teaching and communicating research results.

Collaborative projects are currently underway in non-Euclidean geometries, cosmology, meteorology, and parallel processing. The CAVE is proving itself a useful tool for scientific visualization, in keeping with our Laboratory's goal of providing scientists with visualization tools for scientific insight, discovery, and communication.

8. Future Work

Further research efforts will tie the CAVE into high-speed networks and supercomputers. We have interest in adding motion-control platforms and other highly tactile devices. Hardening and simplifying the CAVE's design for the nation's science museums, schools, and shopping malls is a goal as well. Design and implementation of quantitative experiments to measure CAVE performance are also planned.

9. References

- [1] Bishop, G., Fuchs, H., et al. Research Directions in Virtual Environments. *Computer Graphics*, Vol. 26, 3, Aug. 1992, pp. 153-177.
- [2] Brooks, F.P. Grasping Reality Through Illusion: Interactive Graphics serving Science. *Proc. SIGCHI '88*, May 1988, pp. 1-11.
- [3] Cruz-Neira, C., Sandin, D.J., DeFanti, T.A., Kenyon, R., and Hart, J.C. The CAVE, Audio Visual Experience Automatic Virtual Environment. *Communications of the ACM*, June 1992, pp. 64-72.
- [4] Codella, C., Jalili, R., Koved, L., Lewis, B., Ling, D.T., Lipscomb, J.S., Rabenhorst, D., Wang, C.P., Norton, A., Sweeny, P., and Turk, G. Interactive simulation in a multi-person virtual world. *ACM Human Factors in Computing Systems, CHI '92 Conf.*, May 1992, pp. 329-334.
- [5] Chung, J.C., Harris et al. Exploring Virtual Worlds with Head-Mounted Displays. *Proc. SPIE*, Vol. 1083-05, Feb. 1990, pp. 42-52.
- [6] Deering, M. High Resolution Virtual Reality. *Computer Graphics*, Vol. 26, 2, July 1992, pp. 195-201.
- [7] Fisher, S. The AMES Virtual Environment Workstation (VIEW). *SIGGRAPH '89, Course #29 Notes*, Aug. 1989.
- [8] Max, N. SIGGRAPH'84 Call for Omnimax Films. *Computer Graphics*, Vol 16, 4, Dec. 1982, pp. 208-214.
- [9] McDowall, I.E., Bolas, M., Pieper, S., Fisher, S.S. and Humphries, J. Implementation and Integration of a Counterbalanced CRT-based Stereoscopic Display for Interactive Viewpoint Control in Virtual Environments Applications. *Proc. SPIE*, Vol. 1256-16.
- [10] Plato. *The Republic*. The Academy, Athens, c.375 BC.
- [11] Rheingold, H. *Virtual Reality*. Summit, New York, 1991.
- [12] Sutherland, I.E. The Ultimate Display. *Proc. IFIP 65, 2*, pp. 506-508, 582-583.
- [13] Teitel, M.A. The Eyephone: A Head-Mounted Stereo Display. *Proc. SPIE*, Vol. 1256-20, Feb. 1990, pp. 168-171.
- [14] *Graphics Library Programming Guide*. Silicon Graphics, Inc. 1991.

Acknowledgments

CAVE research is being conducted by the Electronic Visualization Laboratory of the University of Illinois at Chicago, with extraordinary support from Argonne National Laboratory and the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign. Equipment support is provided by Ascension Technology Company, DataDisplay Corporation, Electrohome Projection Systems, Polhemus, Silicon Graphics Computer Systems, Stereographics Corporation, and Systran Corporation. Major funding is provided by the National Science Foundation (NSF) grant ASC-92113813, which includes support from the Defense Advanced Research Projects Agency and the National Institute for Mental Health, NSF grant IRI-9213822, and the Illinois Technology Challenge Grant.