Visual Requirements for Virtual Environment Generation

Robert V. Kenyon, Thomas A. DeFanti,
Electronic Visualization Laboratory
Department of Electrical Engineering and Computer Science
The University of Illinois at Chicago, Chicago, IL
Daniel J. Sandin
Electronic Visualization Laboratory, School of Art and Design
The University of Illinois at Chicago, Chicago, IL

Abstract

This paper describes how a requirement for stereovision can impact virtual environment characteristics and performance. Human visual anomalies that result from the limitations of producing congruent visual cues are described along with possible solutions. The CAVE is used as a model environment that implements stereovision to provide users with visual contact with objects at arms' length.

Introduction

Virtual Environments (VE) enjoy such a wide range of applications that defining a single set of visual requirements for image generation is impossible. For example, flight simulation is a reasonably mature VE application whose goals and performance measures are well established. Nevertheless, one must first understand how the environment will be used before specifying the visual requirements for this VE application. Unfortunately, there is no universal solution to such a problem: "the devil is in the details."

This paper will focus on a class of VE applications where the user directly manipulates virtual objects within arms' length. Such applications require the use of stereovision to render a believable environment, and thus produce an immersive experience for the user. The effects of stereovision on system performance and human interaction will be examined for the CAVE Automatic Virtual Environment (CAVE).

The CAVE Automatic Virtual Environment (CAVE)

The CAVE is a projection-based virtual environment system that surrounds the viewer with up to four screens (Figure 1) and allows both physical and virtual objects to occupy the same space (1). The screens are arranged in a 10 foot cube made up of to two or three rear-projection screens for walls and a down-projection screen for the floor. Each (Electrohome Marque 8000) projector's optics is folded by mirrors due to room size limitations. The images projected onto the CAVE walls are controlled by an SGI Onyx with three Reality Engine 2s. Each Reality Engine is dedicated to rendering the images for one wall of the CAVE limiting the current configuration to two walls (front and left) and the floor, unless another Onyx is used.

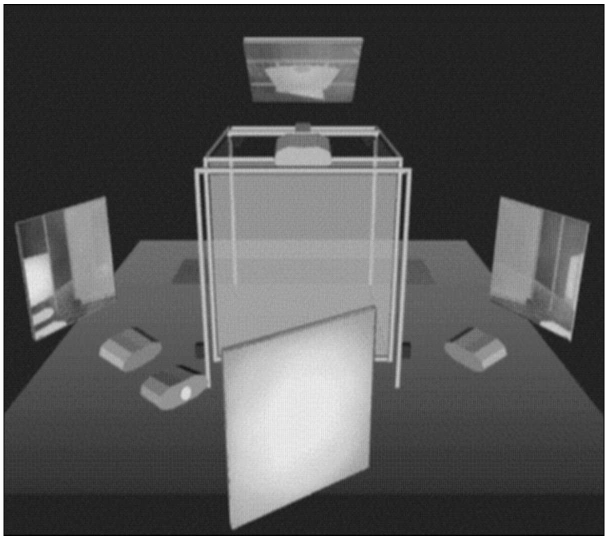

Figure 1: A rendering of the CAVE Automatic Virtual Environment (CAVE) showing projectors and folded optics providing four active surfaces surrounding the user.

The CAVE has an inside-out viewing paradigm where the design is such that the viewer is inside looking out as opposed to the outside looking in. The CAVE uses “window” projection where the projection plane and the center of projection relative to the plane are specified for each eye, thus creating an off-axis perspective projection (2). The correct perspective and stereo projections are based on values returned by the Ascension position sensor attached to the Stereographics Crystal Eyes stereo shutter glasses. Each screen updates at 96 Hz or 120 Hz with a resolution of 1025x768 or 1280x492 pixels per screen, respectively. Two off-axis stereo projections are displayed on each wall. To give the illusion of 3-D, the viewer wears stereo shutter glasses that enable a different image to be displayed to each eye by synchronizing the rate of alternating shutter openings to the screen update rate. When generating a stereo image, the screen update rate is effectively cut in half due to the necessity of displaying two images for one 3-D image. Thus, with a 96 Hz screen update rate, the total image has a maximum screen update rate of 48 Hz. The CAVE has a panoramic view that varies from 90° to greater than 180° depending upon the distance of the viewer from the projection screens. The direct viewing field of view is about 100° and is a function of the frame design for the stereo glasses.

Head and hand position are measured with the Ascension Flock of Birds six degree-of-freedom electromagnetic tracker operating at a 60 Hz sampling frequency for a dual sensor configuration. The transmitter is located above the CAVE in the center and has a useful operating range of 6 feet. Head position is used to locate the eyes to perform the correct stereo calculations for the observer. The CAVE's second position sensor is used to allow the viewer to interact with the virtual environment. Since this system is nonlinear and such nonlinearities can significantly compromise the virtual experience of immersion for the user, a calibration of the tracker system is needed. Nonlinearities caused by the metallic objects and electromagnetic fields created by other devices resident in and about the CAVE are compensated to within 1.5% by linearizing values returned by the head tracking system using a correction table containing calibrated positions in the CAVE (3).

Stereovision Requirement and Implications

Current VE applications are in some ways more ambitious and run on systems that have less computational power than current flight simulation applications. A chief attribute provided by most VE applications that can impact system performance, is user interaction with proximal virtual objects. To work effectively with objects at close range a user requires that the VE provide stereovision. This one necessity alone creates a series of constraints affecting the virtual environment. Stereovision requires that the user’s current head position and orientation in the space be used so that the correct perspective views for each eye are generated. Without such information the 3-D world appears distorted. Consequently, the need to know head location, forces the use of head tracking equipment that can compromise overall system performance in areas such as image update rate and lag.

Stereovision

The introduction of stereovision in the CAVE, for instance, produced many advantages and some drawbacks. Clearly, adding stereo allows a more natural processing of near object cues from the visual environment. Size, distance, location and navigation to the objects become more natural and less ambiguous to the user. However, there are performance penalties to be paid. To supply the two independent view points to the user, the resolution of the display is reduced (using the Crystal Eyes shutter glasses) since the same video display is used to present the right and then the left eye information using field sequential video. This has the effect of reducing resolution by 26% from 1024 horizontal lines to 748 lines. In addition, to prevent 30 Hz flickering of the stereo image that usually accompanies a single frame refresh rate of 60 Hz, the refresh rate must be increased in excess of 96 Hz for a virtually flicker free stereo rate of 48 Hz for the highest vertical resolution. So instead of a maximum update rate of 60 Hz we must settle for a 20% reduction to 48 Hz.

However, the reduction in resolution and update rate could be overcome with some design changes to the CAVE display system. For example, doubling the number of projectors per screen along with the number of graphics processors would restore the display to the original resolution (1024 horizontal lines) and update rate (96 Hz). To restore stereovision without shutter glasses the user would wear passive crossed polarizers with matching polarizers on the corresponding projector.

Viewer Centered Perspective

Another penalty in performance occurs in generating *two* off-axis perspective projection images slaved to the head for each update. In most vehicle simulation environments, the operator's perspective view is fix to the heading of the vehicle and not to the operator's direction of gaze. For example in most flight simulators, the eye point used for the visual perspective, while located close to the expected location of the pilot's head, is fixed to the axis of the vehicle and not to the pilot's head direction. This situation results in only one correct viewing direction for the rendering of the visual scene. Movements of the head and eyes to locations away from the direction-of-projection can result in a somewhat distorted perspective view for the operator. However we should note that in dome simulators, the distance from the center of projection (pilot's eye) to the projection plane (walls of the dome ~ 10 ft.) is large enough, that the head motion in the cockpit has a negligible effect on the overall correct perspective seen by the pilot.

In the CAVE and other head-tracked virtual environments, the perspective view is generated using the direction-of-projection determined by the measured position and orientation of the operator's head. Without this feature the farther the user is from the true center of projection the more distorted the image of near objects appear.

The need to track the user's head for this and other good reasons can add a great penalty to the performance of the system. The generation of the images is now at the mercy of the head tracking instrument's performance. These systems can add long and some cases unacceptable delays between user motion and the resultant motion on the screen. This is especially true for magnetic systems which have gained much popularity since they give the operator freedom to roam about the environment. In addition to the lag, these systems are nonlinear near the edges of the tracker range. These nonlinear errors can so distort the image that objects can appear to fly away from the observer as they are approached. To counteract these effects and make this environment useful for training physical world tasks, calibration of the tracker within the working space is needed (3). This can add more complexity to these systems and more computations per image.

Co-existing Physical and Virtual Objects

One of the advantages that the CAVE affords its users is the ability to see both physical and virtual objects simultaneously. This permits the user to directly view his/her own body, limbs, hands, or that of another person in the CAVE. Consequently, we do not need to spend energy on modeling or rendering poor replicas of the real thing. This not only improves graphics performance but also manual task performance as well (4). The size and location of physical and virtual objects introduce the ability to now interact with objects from both worlds making the environment more powerful tool for prototyping and for realistic interactions with objects in the environment. This advantage also introduces anomalies into the visual world viewed by the user. For example, physical objects can occlude virtual objects but the reverse is not true. In addition, conflict between accommodative and convergence stimuli furnished by adjacent physical and virtual objects within the work space can lead to eye strain and visibility problems within the environment.

Oculomotor Interactions

In the physical world we change our gaze to objects at different distances from us without much notice of what occurs within the motor control portion of the visual system. As we bounce our gaze from object to object, our oculomotor system changes the focus (accommodation) and direction of gaze for the eyes (convergence) so that the object of interest is clear and single. In current VE systems, all virtual objects produce the same accommodative stimulus to the eye regardless of their perceived distance by the user.

In cases where all the objects are far from the eye (> 2 meters) there is little concern for such discrepancies. However, these two systems are placed in conflict with each other when physical and virtual objects exist simultaneously within a VE's working space. The resulting image anomalies can produce disturbing visual consequences for the user.

In 3-D projection based systems like the CAVE and for see-through VE systems, interactions with concurrent physical and virtual objects are a natural consequence of their design. However, the optical characteristics of the physical and virtual objects can be very different. A physical object a foot away from you produces convergence, accommodative, and stereo stimuli that are all in accord. Virtual objects can produce congruent convergence and stereo stimuli for an object a foot away, but the accommodative stimulus is determined by the optical system of the VE. In the case of the CAVE, the accommodative stimulus of a virtual object is determined by the viewer's distance to the screen.

If only other virtual objects are used to interact with the VE object, the visual conditions are not immediately a problem. When you introduce a physical object near the same location the large difference in the stimulus to accommodation between each object can immediately result in one of the two objects becoming blurred and doubled. For example, while sitting in real chair at a virtual table you pickup a virtual soda can on the table with your real hand, the can's optical distance is at the screen while your hand is at arm's length. If you attended to the can, your hand would be out of focus and perhaps double while the can would be sharp and single. If now you attended to your hand, the can would be out of focus and doubled. This dual forced-choice condition placed on the user when interacting with both physical and virtual objects may limit how we apply this technology.

Prospects for training users to differentially accommodate so that both objects can be seen clearly are not good. However, many people can learn to see both near and far objects simultaneously in focus when one eye wears a distance correction and a far correction for the other (monovision). Schor et. al. (5) has found that presbyopic contact lens patients with monocular corrections could see clearly at all distances by suppressing blur that occurs regionally between corresponding retinal areas. This interocular suppression of blur over corresponding areas of the retina is believed to be the main mechanism for the successful use of monovision. In our situation, a monocular corrective lens worn by the user might prove to be a solution to the anomaly produced by co-existing physical and virtual objects. The prospects for successfully using this method in the CAVE is uncertain since image brightness is low in the CAVE. However, we plan to evaluate the utility of this technique in the CAVE before long.

Conclusions

The demand for more realism in VE has historically pushed the technology to its limits and has encouraged new ways to produce real-time imagery. In the example of stereovision, low cost and effective means of providing stereo from desktop to room sized VE systems have expanded the scope of VE applications. In exchange, we have sacrificed visual resolution and maximum update rate, endured increases in image lag, nonlinear image motion and perceptual anomalies. Despite this sacrifice, the visual cues generated in these environments are incomplete and some times inconsistent with the physical world. Nevertheless, stereo cues have proved to be immensely useful to the human operator in VE. The utility of these cues despite all these problems is yet another testament to the adaptability of the human to the environment. Without this kind of built-in servo system, it is unlikely that virtual environments would be as useful as they are today.

References

1. Cruz-Neira C, D.J. Sandin, and T.A. Defanti. Surround-Screen Projection Based Virtual Reality: The Design and Implementation of the CAVE. Computer Graphics, 27: 135-142, 1993
2. Bos, P.J. Performance limits of stereoscopic viewing systems using active and passive glasses. Proceedings of the IEEE Annual Virtual Reality International Symposium (VRAIS) (Seattle, WA, Sept. 18-21), 371-376, 1993.
3. Ghazisaedy, M., D. Adamczyk, D.J. Sandin, R.V. Kenyon, T.A. Defanti. Ultrasonic Calibration of a Magnetic Tracker in a Virtual Reality Space, Proceedings of the IEEE Annual Virtual Reality International Symposium (VRAIS) (Raleigh, NC, March 11-15), 179-188, 1995.
4. Kenyon R and Afenya M, Training in Virtual and Real Environments, Annals of Biomedical Engineering (accepted).
5. Schor, C., Landsman, L., Erickson, P., Ocular dominance and the interocular suppression of blur in monovision. Am J. Opt. & Physiol. Optics, 64: 723-730, 1987.