Interactive stereoscopy optimization for head-mounted displays

Patrick Min and Hans Jense

TNO Physics and Electronics Laboratory
P.O. Box 96864
2509 JG The Hague
The Netherlands

ABSTRACT

In current Virtual Environment systems, the stereoscopic images presented in a Head-Mounted Display are far from optimal. The aim is to achieve orthostereoscopy, which roughly means images should "behave as in real life". A theoretical model of stereoscopic optics was used to implement a test and optimization system. Tests were devised to analyze the importance of many stereoscopy related parameters. The system's capability to independently set these parameters allows an optimization of the stereoscopic images, given the limitations of the display device used.

1. INTRODUCTION

Especially in more serious virtual environment (VE) applications such as simulation and training systems it is essential that the virtual environment mimics its real counterpart as much as possible. This includes a realistic three-dimensional image of the environment. Ideally, the image is orthostereoscopic, or as Sutherland states:

> The image presented by the three-dimensional display must change in exactly the way that the image of a real object would change for similar motions of the user's head.[8]

In most current VE systems stereo images are produced without taking all of the factors that influence the stereoscopic quality into account. Instead of using an accurate computational model of the Head-Mounted Display (HMD), most HMD parameters are simply ignored, or set to average values. This results in an image that is generally not orthostereoscopic, i.e. not three-dimensionally realistic, and sometimes even causes eyestrain. In section 2 we discuss the errors that may occur for each parameter.

Robinett and Rolland have presented a computational model for the stereoscopic optics in an HMD.[7] This model was used as the basis for the implementation of a system allowing the independent manipulation of stereoscopy related parameters. A review of this model is given in section 3.

In section 4 an overview is provided of the system that was developed based on the hteoretical model. Both the hardware used and the software functionality are described. Tests were devised to optimize various parameters. Twelve people were tested in order to evaluate the effects of changes in the parameters. The results are presented and discussed. Finally, in a concluding section, suggestions on how to achieve better stereoscopy, and for future research are given.

2. POSSIBLE ERRORS

The errors that may occur in the computation of stereo image pairs can be classified in two categories:

1. general, not HMD specific errors
2. HMD specific errors

Both types of errors will be discussed in more detail.

0-8194-1472-7/94/$6.00

2.1. General errors

So called general errors result from an incorrect simulation of reality, or more precisely of the way our eyes see in real life. This can either involve elements of the physiology of the human eye or the mathematics behind the image projection in the eyes.

2.1.1. Accomodation/convergence errors

Eyes are accustomed to convergence and accomodation being correlated: for every distance there is an appropriate convergence angle (such that the eyes are turned towards an object at that distance,) and accomodation (to bring the object into focus). A display screen normally is positioned at a fixed distance, hence the eyes have a constant accomodation. But the convergence of the eyes corresponds to the apparent distance of an object in the virtual environment. Veron calls this phenomenon an accomodation/convergence conflict.[9] As a remedy, Robinett and Rolland suggest the user must learn to decouple accomodation and convergence.[7]

2.1.2. Incorrect Inter-Pupillary Distance

The inter-pupillary distance (IPD) of a viewer determines how much the eyes must converge in order to focus on an object at a specific distance. If an average IPD is assumed in the rendering computations (e.g. 65 mm), a viewer with a larger IPD would perceive the object at too large a distance, while someone with a smaller IPD would think the opposite.

2.1.3. Incorrect projection

The projection mathematics involved in the computation of stereoscopic image pairs are usually based on some approximation of the real-world situation. Some frequently used projection techniques (in increasing order of accuracy) are briefly described below:

Projection along parallel viewlines For efficiency reasons, the perspective projection may be implemented assumiong parallel viewlines. in this case the convergence angle never corresponds with reality, except when focus is at infinity (for instance, when looking at a star.)

Rotation A method sometimes used to generate the left- and right-eye views is to rotate the 3D scene by a few degrees both counter-clockwise to obtain the left-eye image, and clockwise to obtain the right-eye image respectively. This method ususally introduces vertical parallax (displacements) in the images, which may cause severe eyestrain.[1]

On-axis projection On-axis projection uses parallel viewlines and one center of projection. The required horizontal shift for the left- and right-eyes are obtained by translating the object. Hodges describes an algorithm for on-axis projection.[1] Roughly, the algorithm works as follows:

for the right eye view:
- translate the object data to the left by IPD/2
- perform the standard perspective projection
- pan the resulting image back

For the left eye view the translation and pan are performed in the oppisite directions. The field of view of on-axis projection is the same as for a single perspective projection. Williams and Parrish show that for example for a 40 degrees horizontal FOV per eye the binocular FOV is 35 % smaller than it would have been if an off-axis projection (discussed next) had been used.[10]

Off-axis projection The off-axis projection assumes converging viewlines and two centers of projection, one for each eye.[1] It most closely corresponds with reality, because the eyes now converge towards the object in focus. The only problem here is to find out on which object the viewer is focusing, because this determines the convergence angle.

2.2. HMD specific errors

The role of the HMD in a typical VE system is shown in figure 1. Errors which result from not, or incorrectly, incorporating the properties of the HMD itself in the rendering calculations, are considered in this section.

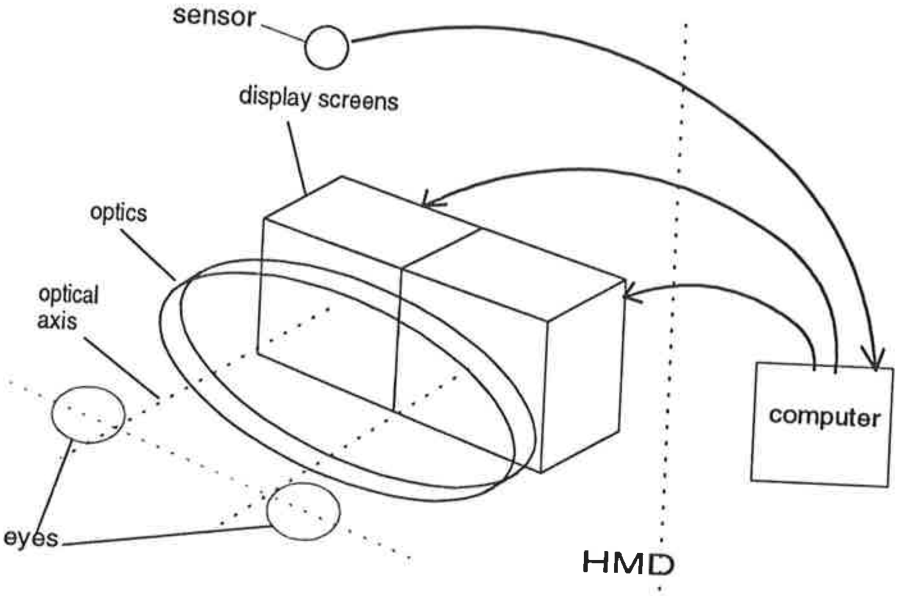

Fig. 1. The HMD in a virtual environment system.

2.2.1. Positional errors

If the optical axes were parallel, and passed through the center pixels of the screens and through the centers of the eyes, turning on the center pixels would show a dot positioned at infinity. But the axes may not be parallel, the screen centers may be offset from the axes, and the eyes may be offset with respect to the axes as well. Errors resulting from these deficiencies can be called "positional errors". Some typical examples are:

Failure to account for angle between optical axes When the optical axes are not parallel (e.g., due to manufacturing tolerances/restrictions), this has to be corrected by a rotation of the left and right-eye image, such that it balances out the rotation.

Failure to incorporate position of screens If the screen centers are offset from the optical axes, all displayed data are offset. In case of a horizontal offset, the eyes need a different (incorrect) convergence angle to focus on an object. A vertical offset results in a height error.

Failure to incorporate Inter-Pupillary Distance In addition to using the correct IPD in the projection, it is also important with respect to the HMD. If the viewer has an IPD equal to the distance between the optical axes, the images are positioned correctly: both centers of projection are located exactly in front of their respective eyes. If the IPD differs from this distance, the images are at a horizontally incorrect position, resulting in a convergence error. In an HMD with mechanical IPD adjustment this problem does not occur, as the screens themselves are moved to get the centers positioned right.

Incorrect Field Of View The Field of View (FOV) used in the projection computations should be the same as the FOV actually experienced by the viewer, i.e. the FOV actually subtended by the images of the display screens. If the computational FOV is too small, then the displayed object will appear too large, and vice versa.

2.2.2. Optics errors

The optics used to project the images from the display screens in the viewer's eyes are of course never perfect. While we do not intend to enter in an exhaustive discussion of all kinds of optical defects and aberrations, two specific errors need to be addressed here from a practical point of view.

Non-linear distortion When a wide Field Of View is projected (warped) onto a flat plane, distortion is inevitable.[2] The LEEP optics used in many HMDs use a fish-eye like transformation to achieve this. This means that the largest part of the image area is devoted to the central part of the FOV, and that the peripheral area is compressed into the side of the image. (Fortunately this distortion corresponds with the relative importance of the various parts of the human FOV.)

When a flat plane is seen through such optics, the magnification is larger for points that are further from the optical axis. This is called a positive or pin-cushion distortion, and causes lines that are straight on the display screen to be curved in the virtual image. In the next section a model for this distortion is discussed, as well as an approximate inverse distortion to correct the error (called a negative or barrel distortion).

Chromatic aberration Differently coloured light rays diffract differently in the lens system, causing lateral chromatism, or "chromatic difference of magnification".[2] In the LEEP optics, blue is magnified about 1% more than red, with green in between. This error is especially noticable in the peripheral part of the FOV.

3. THE COMPUTATIONAL MODEL

In order to compute correct projections of the 3D image space, several HMD specific parameters need to be incorporated in the calculations. A computational model for the optics in an HMD given by Robinett and Rolland will aid in determining these parameters.[7] The general ideas involved are reviewed in this section. For a thorough discussion of the model the reader is referred to the original article.

3.1. Model for one eye

The model relates the radial position of a pixel on a display screen inside the HMD (r_s) with to the radial position of its corresponding pixel as perceived on the virtual image (r_v). Both positions can be normalized by dividing them by their maximum possible values, yielding r_{sn} and an r_{vn}. If the optics had no distortion, then $r_{vn} = r_{sn}$. But they have, so an (approximate) correction term must be added: $r_{vn} = r_{sn} + k_{vs}r_{sn}^3$. The coefficient k_{vs} is a measure for the amount of distortion.

To remove the distortion, an image should be transformed by the inverse transformation. This is called predistortion. So to be able to predistort an image, we need to find the inverse of the function above. An exact closed-form expression is not possible, so an approximation is used again: $r_{sn} = r_{vn} + k_{sv}r_{vn}^3$. Robinett and Rolland show that this approximation is at worst about 2 % off from the correct value. (This is an estimated error, measured from a graph in the article.)

3.2. Model for two eyes

So far we considered just one eye. We need to extend the model to include two eyes in order to:

1. calculate a correct FOV
2. incorporate the offset of the screen centers from the optical axes

These issues are described in the following sections.

3.2.1. Calculation of a correct FOV

The simplest way would be to assume that the optics have linear magnification, but this obviously results in an incorrect computational FOV. We should account for the distortion. First we compute the radial positions in the virtual image of the top, bottom, left and right side of a screen. Using the distance of the virtual image we then get the angular position of each side, and consequently the horizontal and vertical FOV.

Another method is analytical ray tracing: from the exact optics specifications (for each lens in the lens system) the exact path of a lightray passing through each lens surface is calculated. This yields a slightly more accurate FOV than the previous method.

3.2.2. Incorporation of the offset of the screen centers from the optical axes

Finally a correction is needed if the center of a display screen is offset from the optical axis. To correct this error, a perspective projection is necessary that has its computational center of projection at that offset. Another method is to calculate the offset in pixels, and translate the screen image by that offset in the opposite direction. Note that in this case the extent of the screen areas must be larger than the computed screen images, or else data will be lost when the images are translated.

4. THE TEST SYSTEM

In order to determine the relative importance of the errors described in the previous section, a test system was developed that allows independent manipulation of many of the parameters in the rendering of stereoscopic images. This section describes the hardware setup and the software functionality of the system.

4.1. Hardware

The equipment consisted of the following (off-the-shelf available) items:

- computer: a Silicon Graphics 4D240VGX graphics workstation with a VideoSplitter, enabling output of up to four arbitrary quadrants of the operator screen
- monitor: a colour monitor with a screen resolution of 1280 by 1024 pixels
- converter: this device converts two RGB video signals with NTSC (RS170A) timing, produced by the VideoSplitter, to the composite NTSC signals required by the HMD
- HMD: a Virtual Research Flight Helmet, containing LEEP optics and two LCD screens with 320 by 200 LCD pixel resolution. Each pixel is part of a "colour triad" and is either red, green or blue, giving a colour resolution of approximately 185 by 140 pixels. The total (binocular) Field Of View is approximately 100 degrees horizontally and 60 degrees vertically.

Figure 2 shows the hardware setup. The operator controls the setting of all parameters from the console. The resulting images are observed by someone (the viewer) wearing the HMD.

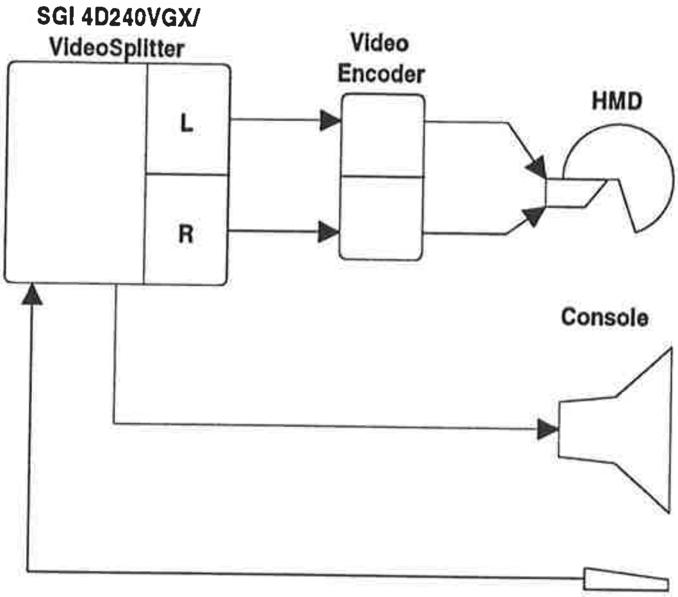

Fig. 2. Hardware setup of the test system.

4.2. Software

The software of the test system was written in C++. The graphics routines use the Silicon Graphics GL library, and the user interface routines are based on the Forms Library.[5]

The operator screen displays the images sent to each HMD screen, a menu and a status window show the values of most parameters. An option is available that allows the viewer to switch between viewing the 3D image or the menu and status window inside the HMD, to enable one person to operate the entire system. In practice, this only makes sense when the resolution of the HMD screens is sufficiently high.

4.3. Error Corrections

While not all of the errors mentioned in the previous section can be corrected with our system, methods for those that can are described below.

4.3.1. Projection types

The available projections are the on- and off-axis projections as well as projection using parallel viewlines. Because of the immediately obvious shortcomings in the method, generating stereoscopic images by means of rotation of the 3D scene has not been implemented.

4.3.2. Compensating for non-parallel optical axes

If the optical axes are not parallel, the left and right image are rotated clockwise and counter-clockwise respectively (through the corresponding eyepositions), both by half the angle between the axes.

4.3.3. Incorporating the correct FOV

From the optics specification and the position of the screens w.r.t. the optics the horizontal and vertical FOV are calculated. It is also possible to interactively change the horizontal and vertical FOV.

4.3.4. Adjusting the center of projection

Whenever a parameter affecting the horizontal position of the images changes (such as the IPD, or the distance betw the optical axes), the computational center of projection has to be moved. In our system this is done by computing displaying images that are larger than the ones seen in the HMD. Only a part of each image is seen by the viewer. T means that if the center of projection has to be moved, we simply move the position of that part o the image that are s to the HMD.

4.3.5. Predistortion

In order to correct the distortion caused by the fish-eye like projection of the LEEP optics, the corresponding inver distortion has been implemented in the following manner. There is a finite number of pixels that have to be (inversel predistorted, i.e., moved to another location. Hence all destination coordinates can be precomputed and stored in a tab We use a table for both the left and right images, as the optical axis is in a different position in each screen.

4.3.6. Chromatic aberration

Correction of the chromatic aberration has not been implemented. However, one obvious way to do this could be to rend the red, green and blue components of each object in three separate frame buffers, scale these by the correct amount compensate for the aberration, and then combine them.

5. THE TESTS

To assess the system's usefulness for conducting stereoscopic viewing tests and the effects of the error corrections, a tes procedure was designed and twelve persons were subjected to it. The test procedure involved, among others, the followin steps:

1. stereoscopic viewing test
2. physical measurement of the IPD
3. IPD test in a virtual environment
4. predistortion test using a regular grid

5.1. Stereoscopic viewing test

Subjects should be able to view stereoscopically. The test we used is the standard TNO test for stereoscopic vision.[3] The subject wears a pair of red/green colored glasses. A series of random dot stereograms is presented, containing pictures requiring a certain stereo acuity in order to be seen.

5.2. IPD measurement and test

First the subject's physical IPD was measured. Then the subject put on the HMD, in which a special test object was displayed, as shown in Figure 3.

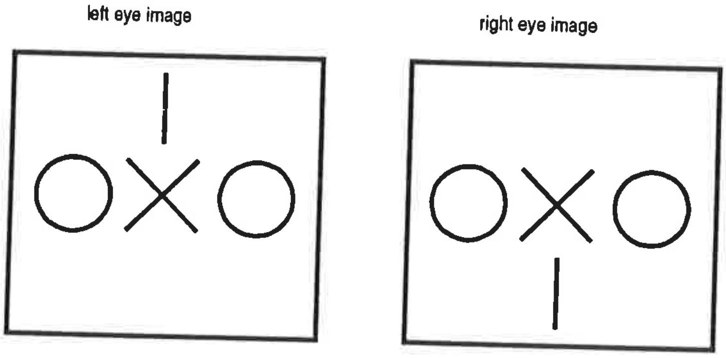

Fig. 3. IPD test object

The fact that the left and right eye images are different is because the brain is very much able to correct for stereo images based on an incorrect IPD. The idea with this test object is that the brain will attempt to fuse the central part of the images (the "OXO"), and probably succeed if the computational IPD is not too far off, while leaving the vertical lines at their true positions, as they are each conflicting with the data received by the other eye. In other words: we assume that the brain behaves differently for different areas of a perceived image.

Initially the computational IPD was set to the measured IPD plus 10 mm. In this way we were sure that the subject perceives an incorrect image, because he generally cannot diverge his eyes. The subject would then change the computational IPD until the "OXO" fused and could be comfortably viewed. The vertical lines however were usually not aligned by then, which made a further, more precise adjustment of the IPD possible.

5.3. Predistortion test

The subject was positioned directly in front of a regular grid, and was asked if the grid lines at the edges of the image either:

- curved outward
- were straight
- curved inward

The predistortion coefficient was then adjusted until the viewer was convinced that the grid lines appeared (approximately) straight.

6. TEST RESULTS AND DISCUSSION

As stated earlier, twelve people were tested using the procedure described in the previous section. Here we discuss the outcome of this study.

6.1. Stereoscopic viewing test

All twelve persons were able to see stereoscopically. The average stereo acuity was 69" ≈ 1'.

6.2. IPD test

The average measured IPD was 65 mm. The IPD test was conducted twice, once with the test object at a distance of two meters, and once at a distance of 0.5 meters. At the smaller distance the test object would almost completely fill the FOV.

The average IPD that appeared most comfortable was 63 mm with the test object at 2 meters, and 66 mm with the test object at 0.5 meters.

The significance of these results is somewhat reduced because the precise mapping of the images on the operator screen to the display screens inside the HMD is not known. For this the HMD must be disassembled (at the time of writing the manufacturer did not have exact data either). For instance, by looking with the right eye through the left eye optics and vice versa it could be seen that:

- not all of the operator screen images (i.e. the part sent to the HMD) is visible on the HMD screens
- the loss of data is different for each side of a screen
- the loss of data is different for the left and right screen

This implies that all correctional translations are inaccurate: they are calculated in operator screen pixels, assuming a certain number of pixels map on a certain width in millimeters on an HMD screen. Compensating for this error (by what is known as "video overscan" or "image cropping") is treated extensively in a recent report by Rolland and Hopkins.[6]

6.3. Predistortion test

The theoretically optimal predistortion coefficient for the LEEP optics in our HMD is -0.18.[7] The average of the coefficients chosen by our subjects to be optimal was -0.17. This coefficient results in a slightly less (barrel) predistorted grid.

The difference between our value and the theoretically optimal one may be explained by two reasons:

- as has been said, the exact operator screen image to HMD screen mapping is not known, causing pixel positions calculated in the precomputing stage to be slightly off
- before the grid is predistorted, the subject sees a pin-cushion distorted grid. This may influence the subject in such a way that he sees a barrel-distorted grid after predistortion, even if this grid is actually straight.

6.4. Relative importance of the parameters

Using our brief experience with the test system, we can attempt to determine the relative importance of each parameter influencing stereoscopic quality.

Correct convergence This depends on the images' horizontal position, which in turn depends on whether or not the IPD, screen center offset, the distance and angle between the optical axes are incorporated in the calculations.

Projection type We found that on-axis projection is especially disturbing for viewing at close distances. If a VE application has just one object of interest (e.g., a tool) at a small distance, one may choose to render it using converging viewlines (= off-axis projection), and the surroundings using on-axis projection.

Optics distortion The optics distortion may also result in incorrect convergence, especially near the edges of the image. Apart from that it obviously causes the image to be incorrect. The importance of this error also very much depends on the type of VE application.

Currently, the (inverse) predistortion is computationally too expensive to be performed during real-time rendering. In our system, after optimization and parallel implementation, a frame rate of 4 Hz may be achieved (that is using the four 25 MHz MIPS 3000 processors of the SGI 2D240VGX). The predistortion can however easily be implemented in hardware, which should be done if real-time rendering is required.

Field of view As an incorrect FOV only results in a (relatively small) size error, we classify it as the least important error.

Note that the resolution of the HMD display screens also determines whether a certain error correction is useful or not: if a positional error does not cause a shift of at least one pixel, it will not be visible anyway.

Concerning the special IPD test object, it might be interesting to see if after incorporation of the exact operator screen images-to-HMD mapping the test can be used to determine one comfortable IPD for all distances. In our opinion the test object should occupy a considerable part of the FOV, so it does not become too easy for the brain to combine both images.

7. CONCLUSION

From our brief experience with the test system, it has already become obvious that the incorporation of as much knowledge as possible that we have about the display system (i.e. the HMD) pays off: when set to the correct IPD, the system rendered convincing, solid three-dimensional objects. The ability to independently vary parameters influencing stereoscopy has proven to be very useful: we may now determine the quality vs. computational cost trade-off of several optimizations.

Future improvements of our system will include:

- accounting for the video overscan
- experimenting with mixed projection types
- experimenting with "IPD setting" test objects
- improvement of the predistortion performance, either through hardware or software

It follows that we recommend the following setup for a general VE system:

- incorporate all parameters that improve convergence
- use off-axis projection (converging viewlines) at least for nearby (say closer than 3 meters) objects, and on-axis projection for further objects
- use predistortion implemented in hardware

After taking these measures, we will come close to achieving orthostereoscopy.

REFERENCES

1. L. F. Hodges, "Time multiplexed stereoscopic computer graphics," *IEEE Computer Graphics and Applications,* Vol. 12, Nr. 2, pp. 20-30, March 1992.
2. E. M. Howlett, "Wide angle orthostereo," in: *Proc. SPIE Vol 1457, Stereoscopic Displays and Applications 2,"* pp. 210-223, 1991.
3. TNO Institute for Perception, *TNO test for stereoscopic vision,* 9th edition, Laméris Ootech, Groenekan, The Netherlands, 1972.
4. P. Min, "Stereoscopy optimization for head-mounted displays," *TNO-report FEL-93-S216,* TNO Physics and Electronics Laboratory, The Hague, The Netherlands, July 1992.
5. M. H. Overmars, "Forms library, A graphical user interface toolkit for Silicon Graphics workstations," Version 2.1, Utrecht University, The Netherlands, 1992.
6. J. P. Rolland and T. Hopkins, "A method of computational correction for optical distortion in head-mounted displays," *Technical report no. TR93-045,* University of North Carolina at Chapel Hill, September 1993.
7. W. Robinett and J. P. Rolland, "A computational model for the stereoscopic oprics of a head-mounted display," in: *Proc. SPIE Vol 1457, Stereoscopic Displays and Applications 2,"* pp. 140-160, 1991.

8. I. E. Sutherland, "A head-mounted three dimensional display," in: *Proc. Fall Joint Computer Conference*, pp. 757-764, 1968.
9. H. Veron et. al., "Stereoscopic displays for terrain database visualization," in: *Proc. SPIE Vol 1456, Stereoscopic Displays and Applications,"* pp. 124-135, 1990.
10. S. P. Williams and R. V. Parrish, "New computational and control techniques and increased understanding for stereo 3-D display," in: *Proc. SPIE Vol 1456, Stereoscopic Displays and Applications,"* pp. 73-82, 1990.

PROCEEDINGS
SPIE—The International Society for Optical Engineering

Stereoscopic Displays and Virtual Reality Systems

Scott S. Fisher
John O. Merritt
Mark T. Bolas
Chairs/Editors

8–10 February 1994
San Jose, California

Sponsored by
IS&T—The Society for Imaging Science and Technology
SPIE—The International Society for Optical Engineering

Published by
SPIE—The International Society for Optical Engineering

Volume 2177